

Perception, Usage and Productivity of Variable
Morphological Rules: Investigations on the
Italian Subjunctive

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Philologischen Fakultät
der Albert-Ludwigs- Universität
Freiburg i. Br.

vorgelegt von

Carmela Pietropaolo

aus Ariano Irpino
(Italien)

Wintersemester 2020/21

Erstgutachter: Prof. Dr. Rolf Kailuweit

Zweitgutachter: Dr. Alice Julie Blumenthal-Dramé

Vorsitzender des Promotionsausschusses

der Gemeinsamen Kommission

der Philologischen und

der Philosophischen Fakultät: Prof. Dr. Dietmar Neutatz

Datum der Disputation: 09.03.2021

Acknowledgements

This dissertation was written during my PhD at the University of Freiburg with a grant awarded me by the Research Training Group DFG GRK 1624 “Frequency effects in language”. The work I put in could not have propelled me this far but for a number of people, whose support was the engine driving my efforts.

I thank my first supervisor, Rolf Kailuweit, for his constant support, words of encouragement during the hard times, and invaluable feedback which provided clarity, helped me chisel and give a necessary structure to the dense mass of my research interests and ideas. I am very grateful to Alice Blumenthal-Dramé for her insightful feedback and words of encouragement during the last stages of my work.

I had the privilege of working in the vibrant ecosystem of the research training group, which sustained my work, trained me and was a constant source of inspiration providing new impulse and guiding me toward the most interesting and promising avenues of research. Stefan Pfänder was the conductor and I am indebted to him and to my first supervisor for a number of research opportunities.

To Anne Lerche and Lars Konieczny I owe the knowledge of how to set up experimental studies, how to test hypotheses and how to use statistical methods and programming tools for linguistic studies. They were always there for productive discussions on how to operationalize my research questions and for helping me with the code here and there. Without them, I would not have developed a passion for statistics and for coding.

Christian Mair, Peter Auer, Bernd Kortmann, Evelyn Ferstl, Daniel Jakob, Eva Staudinger, David Lorenz, Hannah Davidson and Bert Cornillie were involved at different stages of the project and always had insightful feedback as well as plenty of hints of good books and literature that steered me toward the right direction. Just as important was the help of our group's coordinators Michael Schäfer, Nikolay Hakimov and Agnes Schneider whose dedication was immense and whose work was flawless.

Special thanks go to Verena Haser for introducing me to the fascinating world of psycholinguistics and experimental linguistics.

Parts of this work were presented to the 2017 Workshop on Morphological Processing, to the iClave conference and to subsequent conferences. I am indebted to Adam Albright, João Verissimo, Harald Clahsen, Eva Smolka, Tal Linzen, Charles Yang and Vsevolod Kapatsinski for the hard questions, the long e-mails clarifying parts of their work and the face-to-face discussions from which I immensely benefited.

Thank you to my colleagues and friends: Emiel van den Hoven, Luke Bradley, Udo Rohe, Boniface Nkombong, Bella Diekman, Ana Estrada, Karolina Rudnicka, Laura Terassa, Uliana Schöller, Olga Glanz, Pia Hagen-Wiest. I immensely benefited from the exchange of ideas, the reading groups and all the research activities we shared.

I am especially indebted to Sophia Wulfert and Helena Levy, compassionate friends with whom I had the luck to share my research journey until the very end.

Thank you to Renata Savy, a talented linguist and a passionate, brave person to whom I owe my falling in love with linguistics.

Silvana, Stas, Guglielmo, Rossella P., Donatella, Serena, Gianluca, Eleonora, Maria, Arianna, Rossella A., Giuseppe, Vittorio, Elisa and the friends of the Italian-German Stammtisch, thank you for supporting me, putting a smile on my face, advertising my experiments, helping me with equipment and in any way you could. My life as a PhD student would not have been so happy without you.

I owe everything to my family, Salvatore, Antonietta, Raffaella and my grandparents, nonna Carmela, nonna Ida, nonno Ciccio and nonno Raffaele who are not here to share my toil and my achievements.

The backbone of this dissertation coincides with the spinal cord of one person, who gave me more than I deserve and to whom I will be indebted for life. I dedicate this work to Mino, whose intellect is as pure as a crystal and only matched by the purity of his heart.

Contents

List of abbreviations.....xiv

List of figuresxvi

List of tables..... xvii

1 Introduction 1

1.1 The subjunctive mood in the Romance languages..... 1

1.2 Mood variation in the Romance languages 3

1.3 Investigating the factors that influence mood variation in
speech 7

1.4 Models of morphological learning and productivity:
examining the productivity of the subjunctive as a case
study 11

1.5 Outline of the dissertation 18

**2 Perception and conceptualization of mood
variability in folk linguistics23**

2.1 Different types of “subjunctive errors” and social
meanings associated with them 24

2.2 Conclusion..... 61

**3 Semantic accounts of the subjunctive mood in the
Romance languages.....72**

3.1 The quest for the “core” meaning of the subjunctive 72

3.1.1 Different surface realizations of the subjunctive ...	75
3.1.2 A binary opposition between the subjunctive and the indicative	83
3.1.3 Prototypical subjunctive verbs.....	89
3.1.4 Two distinct types of subjunctive.....	91
3.1.5 Weaknesses of the semantic account.....	107
3.1.6 Conclusion	114
3.2 The subjunctive mood at the interface between syntax and semantics.....	116
3.2.1 Temporal characteristics of the subjunctive.....	117
3.2.2 Subjunctive disjoint reference constraint	126
3.2.3 Complementizer deletion	144
3.2.4 Conclusion	156
4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages	159
4.1 The notion of the subjunctive among ancient grammarians.....	164
4.2 The subjunctive mood in unembedded utterances.....	168
4.2.1 Volitive subjunctive	170
4.2.2 Deliberative subjunctive	178
4.2.3 Concessive subjunctive.....	180
4.2.4 Optative subjunctive.....	185
4.2.5 Potential subjunctive	194
4.2.5.1 The conditional sentence.....	199

4.2.6 Conclusion	207
4.3 The subjunctive mood in subordination.....	208
4.4 The diachronic evolution of the subjunctive: from the main clause to the subordinate clause	212
4.5 Conclusion.....	218
5 A synchronic analysis of mood variability: the parameters that affect mood choice	220
5.1 Theoretical background	223
5.1.1 Grammaticalization	223
5.1.2 Frequency effects in language	231
5.1.2.1 Frequency and productivity.....	233
5.1.2.2 Frequency and obsolescence	235
5.1.2.3 Frequency and the mental representation of complex lexical units.....	236
5.1.2.4 Frequency and the erosion of linguistic structure	241
5.1.3 Conclusion	244
5.1.4 Language variation and change	245
5.2 Corpus-based and experimental research on the subjunctive and on mood variability in the Romance languages: a critical review	251
5.3 Effects of lexicalization on mood variability in Italian: a corpus study	268

5.3.1 Materials for the corpus analysis: The C-ORAL- ROM corpus	270
5.3.2 Method.....	272
5.3.3 Results.....	276
5.3.3.1 The influence of the governor verb on the subjunctive/indicative distribution in speech. .	276
5.3.3.2 The mood distribution among high-frequency verbs: a comparison between the subjunctive and the infinitive mood.....	281
5.3.4 Conclusion.....	284

6 The learning and generalization of variable morphological rules: the case of the subjunctive 286

6.1 Models of morphological learning and productivity	288
6.1.1 The role of frequency in analogy-based models ..	292
6.1.2 Summary	295
6.2 Italian verb morphology	296
6.3 Psycholinguistic studies on Italian verb morphology ...	301
6.4 The minimal generalization learner (MGL) and the formation of Romance verbs	304
6.4.1 Results of the experiments based on the MGL algorithm.....	309
6.5 Experiment on the production of novel infinitives and subjunctives: premises and hypotheses.....	315
6.5.1 Method.....	322

6.5.1.1 Participants.....	322
6.5.1.2 Procedure.....	322
6.5.1.3 Simulation with the MGL model	326
6.5.1.4 Materials.....	327
6.5.1.5 Sentence construction and randomization	331
6.5.1.6 Data analyses.....	333
6.6 Results	336
6.6.1 Production of novel infinitives.....	336
6.6.1.1 1 st -conjugation infinitives.....	336
6.6.1.2 2 nd -conjugation infinitives.....	337
6.6.1.3 3 rd -conjugation infinitives	338
6.6.1.4 4 th -conjugation infinitives.....	339
6.6.2 Production of novel subjunctives.....	340
6.6.2.1 1 st -conjugation subjunctives.....	341
6.6.2.2 Subjunctives of the non-default conjugation classes	341
6.6.3 The role of type and token frequency in morphological productivity	342
6.7 Discussion.....	349
7 Conclusion and general discussion.....	354
7.1 Summary of the results.....	354
7.2 General discussion and outlook.....	368
Appendix A: Corpus analysis	376
Appendix B: Experiment stimuli.....	377

Appendix C: Experimental design	382
Appendix D: Collinearity between predictor variables	400
Appendix E: MGL rule reliabilities	402
References	434
German summary.....	503

List of abbreviations

Ø	NULL OR ZERO ELEMENT
1	FIRST PERSON
2	SECOND PERSON
3	THIRD PERSON
ADJ	ADJECTIVE
ABL	ABLATIVE
ACC	ACCUSATIVE
ADV	ADVERB
COND	CONDITIONAL
CONJ	CONJUGATION
CNJ	CONJUNCTION
DAT	DATIVE
DET	DETERMINER
DIM	DIMINUTIVE
FR	FRENCH
FUT	FUTURE
FP	FUTURE PARTICIPLE
IMP	IMPERATIVE
IMPF	IMPERFECT
IND	INDICATIVE
INF	INFINITIVE
INTERJ	INTERJECTION
IT	ITALIAN
LIT	LITERALLY

LT	LATIN
MASC	MASCULINE
NEG	NEGATION
NOM	NOMINATIVE
OBJ	OBJECT
PART	PARTICIPLE
PL	PLURAL
PRF	PERFECT
PPRF	PLUPERFECT
PP	PERFECT PARTICIPLE
PST	PAST TENSE
REFL	REFLEXIVE
PRO	PRONOUN
PRS	PRESENT
SG	SINGULAR
SP	SPANISH
SJCT	SUBJECT
SUBJ	SUBJUNCTIVE
TH	THEME VOWEL

List of figures

Figure 5.1: The phases of grammaticalization.....225

Figure 5.2: Distribution of subjunctive forms by governor.....278

Figure 5.3: Effect of governor on the probability of occurrence of a subjunctive. Estimates of the logistic regression with 95% confidence intervals.279

Figure 6.1: Mean percentages of infinitive responses with associated 95% confidence intervals obtained with human participants and predicted by the MGL based on various frequency measures (token frequency, type frequency and type*token frequency).....345

Figure 6.2: Mean percentages of subjunctive responses with associated 95% confidence intervals obtained with human participants and predicted by the MGL based on various frequency measures (token frequency, type frequency and type*token frequency).....348

Figure 6.3: Comparison between infinitive and subjunctive productivity rates for each conjugation.....349

List of tables

Table 1.1 *Inflectional paradigm of the Italian subjunctive for the 1st and 2nd conjugation class of verbs* 13

Table 3.1 *Subjunctive and imperative present active conjugation Latin verb amāre “to love”* 142

Table 3.2 *The subjunctive explained by generative theories*..... 157

Table 4.1 *Imperative paradigm of Italian verb parlare “talk”* 175

Table 5.1 *Inflectional paradigm of the Italian subjunctive present according to conjugation classes* 227

Table 5.2 *Distribution of subjunctive and infinitive morphology among high-frequency verbs in C-ORAL-ROM*..... 283

Table 5.3 *Type/token distribution of infinitives in C-ORAL-ROM*... 283

Table 5.4 *Type/token distribution of subjunctives in C-ORAL-ROM* 284

Table 6.1 *Relative distribution of verbs by conjugation classes in Italian* 299

Table 6.2 *A representation of the phonological environments discovered by the MGL* 305

Table 6.3 *Relative distribution of the verb classes in Italian and Portuguese*..... 311

Table 6.4 *Inflectional paradigm of the Italian subjunctive present for each conjugation* 315

Table 6.5 *Type/token distribution of infinitives in C-ORAL-ROM*... 319

Table 6.6 *Type/token distribution of subjunctives in C-ORAL-ROM* 320

Table 6.7 <i>Distribution of subjunctive and infinitive morphology among high-frequency verbs in C-ORAL-ROM</i>	320
Table 6.8 <i>Descriptive statistics of the experiment stimuli</i>	329
Table 6.9 <i>Rules generated by the MGL to derive infinitive forms of novel verbs ‘molgo’ and ‘illipingo’</i>	330
Table 6.10 <i>Coefficients of the MAM predicting 1st conjugation infinitives (-are)</i>	337
Table 6.11 <i>Coefficients of the MAM predicting 2nd conjugation infinitives (-ere)</i>	338
Table 6.12 <i>Coefficients of the MAM predicting 3rd conjugation infinitives (-ére)</i>	339
Table 6.13 <i>Coefficients of the MAM predicting 4th conjugation infinitives (-ire)</i>	340
Table 6.14 <i>Coefficients of the MAM predicting 1st conjugation subjunctive (-i)</i>	341
Table 6.15 <i>Coefficients of the MAM predicting subjunctives of the non-default classes, 2nd, 3rd and 4th conjugations (suffix -a)</i>	342
 Appendix Table 1 <i>Estimates of the generalized linear mixed model that predicts the occurrence of the subjunctive based on the governor</i> ...	376
Appendix Table 2 <i>Stimuli used in the oral production task</i>	378
Appendix Table 3 <i>Example of how two different participants may receive the stimulus simo in experiment presentation 1</i>	384
Appendix Table 4 <i>Example of how two different participants may receive the stimulus simo in experiment presentation 2</i>	385

Appendix Table 5 <i>Experiment presentation 1 as experienced by one participant</i>	388
Appendix Table 6 <i>Experiment presentation 2 as experienced by one participant</i>	394
Appendix Table 7 <i>Coefficients of correlation between MGL rule reliabilities based on type frequency and MGL rule reliabilities based on token frequency for the infinitives of each conjugation</i>	401
Appendix Table 8 <i>Coefficients of correlation between MGL rule reliabilities based on type frequency and MGL rule reliabilities based on token frequency for the subjunctives of each conjugation</i>	401
Appendix Table 9 <i>MGL rules for the formation of infinitives with their reliabilities based on type frequency</i>	402
Appendix Table 10 <i>MGL rules for the formation of subjunctives with their reliabilities based on type frequency</i>	420

1 Introduction

The studies presented in this dissertation investigate the perception, usage and productivity of the subjunctive mood in the Romance languages with a particular focus on the Italian subjunctive. A number of theories that attempt to describe the subjunctive are critically evaluated and tested by means of a corpus analysis. In particular, it is explored whether the occurrence of the mood responds to a semantic principle, e.g. the subjunctive is selected in utterances that encode non-factual, non-declarative modality as suggested by semantic-based accounts (Farkas, 1992b; Giorgi & Pianesi, 1997 among others). Additionally, the productivity of different forms of the subjunctive paradigm is tested by means of an oral production task. The experiment examines how frequency of use impacts the productivity of morphological patterns.

1.1 The subjunctive mood in the Romance languages

The subjunctive is a morphological system inherited from Latin whose main function according to traditional accounts is to encode irrealis, non-declarative, non-factual or subjective utterances. From a functional perspective, it is considered to be in binary opposition to the indicative, the mood of realis, declarative, factual and objective propositions (Bolinger, 1968; Haspelmath, 2003; Hooper, 1975; Hooper & Terrell, 1974). The subjunctive is used in main clauses and primarily in subordinate clauses to convey special modalized meanings, like optative, deontic, epistemic, volitional and evidential interpretations

1 Introduction

(Chafe, 1995; De Mulder & Lamiroy, 2012; Farkas, 1992b; Givón, 1978; Thompson, 1998). It expresses wishes, commands, doubts, counterfactuals etc. The indicative, on the other hand, occurs in assertive/declarative contexts to convey the truth of the event in the actual world at the time of the locution. The language sample taken from Giorgi and Pianesi (1997, pp. 194–200) exemplifies this semantic divide in the Romance languages: (1) shows the indicative mood in main clauses with declarative/affirmative illocutionary force whereas in (2) the subjunctive is selected in unembedded clauses with optative reading. In complement sentences in (3) the subjunctive is governed by volitional predicates.

(1)

- | | | |
|---|------------------------------------|-----------|
| a | Gianni è/*sia arrivato | (Italian) |
| | Gianni has (IND) / (SUBJ) arrived. | |
| b | Jean a/*ait arrivé | (French) |
| | Jean has (IND) / (SUBJ) arrived. | |

(2)

- | | | |
|---|----------------------------|-----------|
| a | (Che) Dio ci aiuti! | (Italian) |
| | (That) God us-help (SUBJ)! | |
| b | (Que) Dieu nous aide! | (French) |
| | (That) God us-help (SUBJ)! | |

1 Introduction

(3)

- a. Paolo vuole che tu venga. (Italian)
Paolo wants that you come (SUBJ).
- b. Paul veut que tu viennes. (French)
Paolo wants that you come (SUBJ).
(Giorgi & Pianesi, 1997, pp. 194–200)

However, the subjunctive occurs first and foremost in subordinate clauses. In traditional theories, the basic semantic opposition is considered useful to explain the distribution of subjunctive and indicative forms in subordination (Farkas, 1992b; Giorgi & Pianesi, 1997). Thus, Farkas (1992b) argues that the subjunctive in embedded clauses is only licensed by lexical elements whose meaning is harmonious with the modal interpretations conveyed by the mood. For example, volitive and optative verbs like “want”, “order”, “request”, “wish” etc. select the subjunctive because they encode irrealis, or non-declarative contexts. Verbs that express certainty and commit the speaker to the truth of the proposition, e.g. “say”, “assert” etc., select the indicative.

1.2 Mood variation in the Romance languages

A closer analysis of speech data, however, reveals that the subjunctive usually alternates with the indicative and the conditional mood in the embedded clauses that would select subjunctive morphology according to the language norm. The selection of another mood form in lieu of a

1 Introduction

subjunctive does not generate an apparent shift in the meaning of the sentence. This phenomenon is called mood variation or mood variability and it has been noticed in the Romance languages (Blücher, 2003; Poplack, 1990; Porto Dapena, 1991 among others). In (4), (5), (6) and (7) are examples of the phenomenon of mood variability in Italian, French, Spanish and Portuguese respectively. The sentences encode illocutionary forces, such as optative, deontic, volitional, epistemic etc., that are traditionally considered the semantic domain of the subjunctive in the Romance languages.

The larger lexical and syntactic context where this mood alternation takes place has been called the “locus of variation”(Poplack et al., 2013, p. 162). In the Italian sample reported in (4) a. and b., for instance, the locus of variation is the embedded verb in the complement clause governed by the verb *volere* ‘want’ (matrix verb or governor verb) in the main predication. In the French sample in (5) a. and b., the subjunctive and the indicative are selected in complement clauses governed by the conditional form of verb *aimer* “like” (*j’aimerais* “I would like”).

1 Introduction

(4)

(Italian)

- a. Questa pagina mi sembrava doverosa e l'ho aggiunta dal momento che voglio che tutti capiscano (IND) l'importanza del software libero.

“This page seemed necessary to me and I created it since I want everybody to understand the importance of free software.”

- b. Il fatto è che voglio che tutti capiscano (SUBJ) come sto dentro e cosa faccio vedere fuori!!

“The fact is that I want everybody to understand how I feel inside and what I let transpire on the outside.”¹

(5)

(French)

- a. Tu sais, tu aimerais mieux qu'ils soient (SUBJ) pas là.

“You know, you'd like it better if they weren't there.”

- b. J'aimerais ça que vous verriez (COND) mon mari. Vous allez voir comme c'était un bel homme.

“I'd like it if you saw my husband. You'll see how handsome he was.”

- c. Tu aimerais qu'il guérisse (IND) ta paralysie, parce que je suis paralysée dans le visage.

¹ The language samples in (4) are taken from the ItWac Web Corpus (Baroni et al., 2009)

1 Introduction

"You would like for him to cure your paralysis, because my face is paralyzed."

(Poplack et al., 2013, p. 141, also available in the *Corpus du français parlé à Ottawa-Hull* [Poplack, 1989])

(6)

(Spanish)

- a. No creo que haya (SUBJ) nadie aquí que no pague la renta.
(SP.073.668)

"I don't think there's anybody here who doesn't pay rent."

- b. No creo que hay (IND) que firmar. (SP.086.555)

"I don't think you have to sign."

(Poplack et al., 2013; also available in *Corpus sociolingüístico de la ciudad de México* [Butragueño & Lastra, 2012])

(7)

(Portuguese)

- a. Eu acredito que vá (SUBJ) sair. (PTG.143.356)

"I believe that it will come out."

- b. Eu acredito que ele devia (IND) ter em torno de setenta anos de idade.
(PTG.099.148)

"I believe that he must be around seventy years old."

(Poplack et al., 2013; also available in *Amostra Linguística do Interior Paulista* [Gonçalves, 2008])

1 Introduction

Generally, as studies in variationist sociolinguistics (Coseriu, 1981; Eckert, 2006, 2012; Labov, 1963, 1972; Sankoff & Thibault, 1981) have illustrated, when two or more forms are available in natural discourse to encode the very same information, their selection can be attributed to the social and geographical background of the speakers, to their identity, group affiliation and to the communicative setting. All these dimensions of variation, however, do not suffice to explain all of the most important facts concerning the use of the subjunctive. Indeed, the alternation between mood forms is attested in the speech of the same individual. In the course of one interaction, the same speaker or speakers from the same social and geographical background, i.e. sisters, neighbors, friends etc. may more or less consciously alternate between the two forms and choose one form over the other seemingly arbitrarily. Poplack (1991, p. 240) terms this phenomenon “inherent variability”.

1.3 Investigating the factors that influence mood variation in speech

It is nowadays a matter of debate whether the mood distribution in speech can be understood by relying on a semantic opposition between the subjunctive and the indicative. While there may be some areas where the mood distribution is semantically motivated (Quer, 2009) and there have been interesting proposals to capture the unique properties of the subjunctive at the interface between syntax and semantics, notably in Farkas (1992b), Bianchi (2001) and Kempchinsky (2009) among others, some scholars have pointed out the

1 Introduction

limits and inconsistencies of traditional semantic approaches which fail to capture actual language practices beyond the stiff norm described in traditional grammars (Montrul, 2007, 2009; Poplack, 1990; Silva-Corvalán, 1994b). Also, these accounts commit the analytic fallacy of attributing the semantic properties to the subjunctive which are actually encapsulated by the linguistic elements that determine its occurrence (Blücher, 2003; Poplack et al., 2018). More specifically, they usually associate the meaning conveyed by the governor verb (also called matrix verb) to the subjunctive. For instance, in sentence (8), grammars talk about an epistemic subjunctive because the verb in the main clause that licenses it, *credere* ‘believe’, is an epistemic verb, in the sense that it expresses the degree of knowledge or the assumption of the subject.

(8) Non credo che Maria sia già a casa.

“I don’t think that Maria (SUBJ) is already home.”

According to a number of studies, mood variability is a product of the diachronic development of the subjunctive which is being slowly supplanted by other mood forms (Montrul, 2007; Poplack et al., 2018; Silva-Corvalán, 1994b among others). They have pointed out that the subjunctive is becoming more and more infrequent and may be close to disappearing in speech, with different rates of decline in usage exhibited in the Romance languages (Carlier et al., 2012; Lamiroy & De Mulder, 2011). Moreover, some scholars assume that, from Latin to the Romance languages, the subjunctive has progressively become

1 Introduction

emancipated from the semantic functions that it originally encoded (Bybee, 2002; Poplack, 1990; Poplack et al., 2018), a process known as *desemanticization* or *semantic bleaching* (Haiman, 1991; Lehmann, 2015). As a result, it is nowadays an empty syntactic marker of subordination whose occurrence is prompted by a small number of governor verbs or by other elements of the larger proposition, like the subordinating conjunction. This phenomenon is called *lexical conditioning* or *lexicalization* (Poplack et al., 2013, 2018; Porto Dapena, 1991). It suggests that the subjunctive has conventionalized in a few more or less formulaic phrases (Bybee et al., 1994; Ellis, 2008; Haspelmath, 1998; Poplack et al., 2018), but is in fact no longer productive. This view is essentially opposed to the accounts provided by the studies within semantics, modal logic, and generative theory. They identify a core underlying function of the subjunctive that is able to explain its surface realizations in speech notwithstanding instances of mood variability. Cases of arbitrariness do not defeat a semantic theory (Farkas, 1992b; Giorgi, 2009; Giorgi & Pianesi, 1997; Kempchinsky, 2009).

The phenomenon of mood variation has been far less investigated in Italian than in French and Spanish, especially if one considers corpus or experimental studies. Indeed, most studies of the Italian subjunctive are conducted within the framework of generativist theory and the minimalist program and do not use large samples of natural linguistic data for their analyses (Bianchi, 2001; Giorgi, 2009; Giorgi & Pianesi, 2004a, 2004b; Kempchinsky, 2009).

1 Introduction

The present work analyzes different aspects of the mood variability in Italian. First, it addresses the question of how it is perceived and conceptualized by the speech community. It dwells on the instances in which the occurrence of other mood forms in lieu of a subjunctive are sanctioned by the speech community presenting different types of “morphological errors” and the metalinguistic discourse that they generate. The analysis is meant to supplement corpus studies by giving an understanding of the social meanings attached to morphological variables and of the social dimensions that impact mood variation (e.g. issues of prestige).

Additionally, the dissertation adopts a historical perspective and investigates the grammar of the subjunctive and its distribution in Classical Latin, in Vulgar Latin and in Common Romance by taking into account a varied sample of texts. Comparisons with the domains of application of the subjunctive in Italian and French are conducted in order to gather indirect clues to the diachronic evolution of the mood. The analysis addresses the hypothesis that “the subjunctive was meaningful in Latin” (Poplack et al., 2018, p. 4) and has since then progressively emancipated from its original function of conveying a modal interpretation (Blücher, 2003; Porto Dapena, 1991).

The sociolinguistic analysis and the historical analysis frame the corpus study on the use of the subjunctive in speech in present-day Italian. The objective of the corpus study is to address the questions raised in the literature concerning the grammar and the functions of the

1 Introduction

subjunctive in subordinate clauses. The following questions guide the analysis:

1. How can the use of the subjunctive be described and understood?
2. Is the morphological variation presented completely random?
 - a. Do semantic factors accurately describe the subjunctive/indicative distribution in speech?
 - b. Is the subjunctive fossilized in a number of ‘subjunctive phrases’ or *schemas*?

The study in particular addresses the hypothesis that only a few governors correlate with the occurrence of the subjunctive (Poplack et al., 2013, 2018; Porto Dapena, 1991) and that the mood distribution cannot be modeled by relying on the lexical semantics of the governors or the modal interpretation of the sentence, contrary to what Farkas (1992b) has suggested.

The corpus study also seeks to give a rough estimate of the productivity of the subjunctive in speech by analyzing its frequency and its lexical richness and comparing it to other moods used as a benchmark.

1.4 Models of morphological learning and productivity: examining the productivity of the subjunctive as a case study

Leaving aside the alternation between the subjunctive and other mood systems, and zooming into the structure of the subjunctive paradigm, there is an additional type of morphological variation within the

1 Introduction

subjunctive system itself. In Italian there is no unique set of rules to derive the subjunctive forms for each tense, but rather variable rules. As is the case in other Romance languages, verbs split into so-called conjugation classes. Each class has its own set of inflectional suffixes. The subjunctive has an allomorphic form for each conjugation. In addition, there are verbs and verb groups whose subjunctive forms are derived by various processes of stem allomorphy, phonological changes in the stem, along with the application of the suffix (e.g. *bere* → *beva* “drink.1/3.SG.PRS.SUBJ”, *fare* → *faccia* “do.1/3.SG.PRS.SUBJ”). Table 1.1 shows the variant morphological paradigm dependent on conjugation class taking the 1st and 2nd conjugation for illustration purposes. Only regular inflection processes, i.e. segmentable and predictable changes to the verb stems by suffixation are displayed for now. In order to highlight the suffixes, they are separated from the stem by a hyphen.

1 Introduction

Table 1.1

Inflectional paradigm of the Italian subjunctive for the 1st and 2nd conjugation class of verbs

Present tense		Imperfect tense	
1 st conjugation	2 nd conjugation	1 st conjugation	2 nd conjugation
<i>pens-are</i>	<i>scend-ere</i>	<i>pens-are</i>	<i>scend-ere</i>
“think”	“descend”	“think”	“descend”
<i>pens-i</i>	<i>scend-a</i>	<i>pens-assi</i>	<i>scend-essi</i>
<i>pens-i</i>	<i>scend-a</i>	<i>pens-assi</i>	<i>scend-essi</i>
<i>pens-i</i>	<i>scend-a</i>	<i>pens-asse</i>	<i>scend-esse</i>
<i>pens-iamo</i>	<i>scend-iamo</i>	<i>pens-assimo</i>	<i>scend-essimo</i>
<i>pens-iate</i>	<i>scend-iate</i>	<i>pens-aste</i>	<i>scend-este</i>
<i>pens-ino</i>	<i>scend-ano</i>	<i>pens-asserò</i>	<i>scend-essero</i>

So far, although the subjunctive mood has received a great deal of scholarly attention from both a diachronic and a synchronic perspective, the analysis has always focused on the subjunctive morphology as a whole, without taking into account the allomorphic variation within the paradigm represented in Table 1.1. Yet, it is known from the studies on morphological productivity and obsolescence that the different forms of a paradigm may be variously affected by decline in usage and attrition (Hopper & Traugott, 2003). The dissertation investigates the productivity rates of the subjunctive and infinitive patterns of each conjugation with the use of an oral production

1 Introduction

experiment: native speakers are asked to derive subjunctive and infinitive forms of new verbs. The infinitive acts as a term of comparison.

In addition, the goal of the study is to address a long-standing debate on the mental representation of morphological knowledge and on the factors that influence morphological productivity. In particular, the following research questions will be explored:

- a) How do speakers generalize morphological patterns to new words?
- b) Is the generalization of both regular and irregular morphological patterns influenced by similarity to existing words?
- c) Does frequency of occurrence of the morphological patterns influence their productivity?

The most prominent theories on language learning do not agree on the answers. A major point of controversy concerns regular productive processes (e.g. English past tense *help* > *helped*) and irregular processes (e.g. *do* > *did*, *go* > *went*)².

² The notion of regularity captures the twofold relationship between the morphemes and the units of meaning on the one hand and between the base and the derived form on the other. Complex words derived through regular morphological processes can be broken down into constituent morphemes (segmentability). Each morpheme is identified as a unit with a specific meaning. For instance, the inflected verb form *added* is made up of the lexical morpheme *add* with meaning “unite, join” and bound morpheme *-ed* with meaning “past tense”. Irregular forms, however, are usually generated by unpredictable changes to the base and they cannot be segmented. There is no

1 Introduction

Dual-mechanism models (Pinker, 1991; Marslen-Wilson & Tyler, 1998) make a fundamental distinction between the way in which regular and irregular morphological processes are treated. Regular forms are generated by rules whereas irregular forms are stored in the mental lexicon. Regular morphological changes are applied to all new words that belong to a syntactic class regardless of the lexical and statistical properties. The extension of regular processes takes place by default and it is not influenced by their frequency in speech (Clahsen, 1995, 1997). Irregular forms, on the other hand, are stored in the lexicon. Irregular morphological changes may be extended to new words that bear a similarity to existing irregulars by means of analogy. Their application is influenced by frequency: the higher the frequency of the irregular pattern, the higher the probability that it will be selected with words that are similar to existing irregulars (Prasada & Pinker, 1993; Clahsen, 1999).

Analogy-based models like connectionist models (Rumelhart & McClelland, 1986; McClelland & Patterson, 2002; Hare et al., 1995; Skousen et al., 2002) and the network model (Bybee, 1995), on the other hand, highlight the role of memory vs. computation in the formation of morphological knowledge. No formal distinction in the

one-to-one relation between form and function in the inflected verb *went*. No single unit with the meaning “past tense” can be extracted from the item, lexical meaning and grammatical meaning are fused together, nor is it possible to establish a relation between the base form *go* and the derived form *went* purely on the basis of phonological similarity (Amenta & Crepaldi, 2012, p. 4).

1 Introduction

acquisition and representation of morphological and lexical knowledge is posited. All word forms, both simplex and complex, that are encountered in the process of language acquisition, are initially memorized. Through an analogical mechanism, the brain establishes connections among the words present in the lexicon based on the principle of phonological and semantic similarity. It then extracts morphological patterns that arise as regularities in the lexicon and generalizes them to new lexical items. Crucially, the experience with language continuously shapes the lexical connections on which the morphological patterns are based. This knowledge is probabilistic: the probability that a morphological pattern is selected to derive a new word depends on both the similarity to the previously learned input and its statistics. When an unknown word is encountered, the higher its similarity to a set of learned items at the phonological and semantic level of representation, the higher the probability that the morphological pattern of that set will be applied (Hahn & Nakisa, 2000).

Within analogy-based models there is disagreement concerning the type of statistics that directly influences the productivity of a morphological pattern. In particular, connectionist models (Daugherty & Seidenberg, 1994; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1986; McClelland & Patterson, 2002) hypothesize that the number of words that take a morphological rule (type frequency) and the absolute frequency of the morphological rule (token frequency) increase the probability that it will be used in speech to derive novel

1 Introduction

forms. Therefore, productive morphological processes are simply discovered by speakers because they are the most frequent and are used with the greatest number of diverse words.

Per contra, according to the network model (Bybee, 1995), only type frequency strengthens the connections between words, on which the regularities are drawn, and ultimately boosts the productivity of a morphological pattern.

The experiment tests the hypotheses of the dual-mechanism model and of analogy-based models by establishing whether there is a fundamental difference in the way regular and irregular subjunctive morphology is applied to derive new word forms. Moreover, the experiment specifically examines the contradictory predictions of the network model and of the connectionist models on the type of statistics that influences the productivity of morphological patterns. This task in particular could not be carried out so far even by the most advanced experiments in Albright (2002a), Albright (2002b) and Veríssimo & Clahsen (2014). This is due to the difficulty of constructing an experiment where the effects of type and token frequency can be distinguished successfully and due to the small language corpora that the studies used to extract their frequency measures (Albright, 2002a).

So far, no study on the generalization of subjunctive morphology to novel words has been conducted. Experiments of this type have only been attempted with the infinitive (Albright, 2002a; Albright, 2002b; Veríssimo & Clahsen, 2014) and the past participle (Say & Clahsen, 2002), and unfortunately they explore each pattern individually. Also,

1 Introduction

they yielded inconsistent evidence: phonological similarity and frequency have been found to play a role in both regular and irregular morphological processes in some studies (Albright, 2002b; Albright, 2002a; Albright & Hayes, 2003 among others) and only in irregular morphological processes in others (Say & Clahsen, 2002; Veríssimo & Clahsen, 2014 among others). Thus, it is unclear from experimental results whether regular and irregular inflections are conceptualized and generalized differently.

The subjunctive offers a rare opportunity to test the hypotheses presented because the proportion of regular and irregular patterns and their frequency of occurrence in speech differs from that of other moods. The irregular and unproductive class is larger for the subjunctive than it is for other moods³. It is large enough to attract new words. Therefore, the subjunctive is particularly well suited to explore possible graded effects of frequency and similarity.

1.5 Outline of the dissertation

The dissertation integrates very different areas of linguistic research, reviewing and critically evaluating conflicting views, to give a comprehensive analysis of the subjunctive that has not yet been provided.

³ Two conjugation classes (the 2nd and 3rd conjugation), which gather irregular and unproductive processes, usually have two distinct inflectional paradigms. The subjunctive paradigm collapses together these two verb classes into one class of verbs.

1 Introduction

Chapter 2 introduces different types of morphological variation connected to the subjunctive. It presents a sociolinguistic analysis, examining how these speech phenomena are perceived by non-linguists and the social and metalinguistic discourse that they generate. It dwells on the aspects of mood variation that receive most attention and on those of which non-linguists remain unaware, thereby establishing the limits of folk perception of mood variability and pinpointing the non-standard linguistic practices that have been normalized by the speech community.

Chapter 3 engages with the most prominent accounts of the subjunctive within the semantic tradition, modal logic, generativist theory and the minimalist program (e.g. Farkas, 1992b; Giorgi & Pianesi, 1997, 2004a, 2004b; Bianchi, 2001; Kempchinsky, 2009). It critically evaluates the fundamental contribution of these studies: their attempt at giving a unitary notion of the subjunctive, at establishing a core underlying meaning of the mood that can account for all its surface functions and its syntactic behavior. The promises and the limits of these proposals are evaluated. In particular, extensive evidence coming from language corpora is provided to show where such theories fail to capture the mood distribution in speech. Also, inconsistencies in the theories are pointed out.

Chapter 4 addresses the hypothesis according to which the subjunctive has progressively lost semantic and grammatical meaning from Latin to the Romance languages, a process known as *desemanticization* (Haiman, 1991, p. 154; Blücher, 2003; Porto

1 Introduction

Dapena, 1991) and viewed as part of grammaticalization (Kuryłowicz, 1965, p. 52; Lehmann, 2015). The implicit assumption is that the subjunctive in Latin had a core modal meaning (see Silva-Corvalán, 1994b; Poplack et al., 2018). This perspective is investigated using the framework and methodologies of historical linguistics. Studies describing the grammar and functions of the subjunctive in Latin are reviewed (e.g. Grandgent, 1907; Handford, 1947; Magni, 2010; Calboli, 1969; Harris, 1974; Basile, 2001; Pinkster, 2015), among them also quantitative analyses based on samples of Classical Latin, Vulgar Latin and Common Romance are presented (e.g. Pinkster, 2015). The use of the subjunctive in Latin is compared to its use in the Romance languages, with a focus on Italian to gather evidence on the restructuring and evolution of the mood in relation to other mood forms, such as the conditional, the imperative and the infinitive. Also, the metalinguistic analysis provided by ancient grammarians (e.g. Macrobius, Cledonius and Priscian) is compared to the analyses discussed in chapter 3.

Chapter 5 reviews the theories of grammaticalization (e.g. Kuryłowicz, 1965, p. 52; Haiman, 1991; Lehmann, 2015) and the studies that elucidate the role of frequency in morphological change and morphological productivity (e.g. Rumelhart & McClelland, 1986; Bybee & Slobin, 1982; Bybee, 1995; Bybee & Thompson, 1997; Pierrehumbert, 2002; Yang, 2016). Empirical studies on the use of the subjunctive in the Romance languages are presented (e.g. Poplack et al., 2013; Poplack et al., 2018; Montrul, 2007; Gudmestad, 2010). Their

1 Introduction

results are critically evaluated in light of the theoretical background. These studies are supplemented by a corpus analysis on mood variation in Italian. The study investigates the influence of the governor on the selection of the subjunctive in embedded clauses. It establishes whether the meaning of the governor can account for the mood distribution as suggested by semantic-based accounts (e.g. Farkas, 1992b). Moreover, an analysis of the productivity of the subjunctive mood in speech is conducted. I examine the overall frequency of subjunctive forms, the number of different verbs that occur with the subjunctive (type frequency) and present a simple measure of lexical richness of the subjunctive (type-token ratio). The analysis additionally focuses on the distribution of subjunctives among high-frequency verbs. The distributional properties of the mood are contrasted with those of the infinitive. The latter, being a productive system extensively studied in the literature on morphological productivity (Albright, 2002b; Albright, 2002a; Veríssimo & Clahsen, 2009, 2014), serves to establish a baseline in order to interpret the data and the patterns that emerge from the analysis.

Chapter 6 presents the theoretical debate concerning how morphological knowledge is acquired and how morphological patterns are extended to new words. In particular, the role of frequency on rule formation and productivity is explored by means of an oral production task with adult native speakers of Italian. The different predictions of analogy-based models (e.g. Rumelhart & McClelland, 1986; Bybee, 1995) and the dual-mechanism model (e.g. Prasada & Pinker, 1993;

1 Introduction

Clahsen, 1997; Clahsen et al. 1997) are tested on the formation of novel subjunctives and novel infinitives. The study takes advantage of the peculiar statistics of the subjunctive to model the influence of type and token frequency on rule generalization.

Chapter 7 summarizes the results and the general contribution of the dissertation. It relates together the findings of the corpus study and those of the experiment providing a general discussion of their relevance and their implications for models of morphological learning and rule productivity. Finally, it engages in a deeper reflection on the research questions that remain to be answered outlining concrete directions for future research.

2 Perception and conceptualization of mood variability in folk linguistics

Sociolinguistic research into the variable use of grammatical and phonological material has shown how the study of grammar cannot transcend the social dimension (Labov, 1972; Bucholtz & Hall, 2004; Bucholtz & Hall, 2005) and the perception of language users in a socio-ecological niche (Niedzielski & Preston, 2010). In particular, “it is important to address the relationship between folk perception and language practices because they are fundamentally linked” (Pietropaolo, 2014, p. 3). Metalinguistic judgments and evaluations, linguistic beliefs about style and prestige influence linguistic behaviors. This is especially true in societies that value linguistic normativity, actively cultivate a standard language ideology and promote a formal instruction of its grammar through the schooling system. Speakers construct their identity through grammar in conformity or opposition to the norm. They may choose to discard it as a result of further societal pressures or of communicative/functional needs. More importantly, from a cognitivist perspective, the metalinguistic reflection coming from language users, for instance in the guise of overt comments about grammar errors, reveals the linguistic phenomena that are subject to conscious analysis. In the case of the mood distribution in Italian, it sheds light on the syntactic and semantic contexts and on the mood forms that appear to be salient, highlighting the locutions where the morphological variation is consciously perceived by the speakers. It thus opens up the possibility to study the meaning that is attributed to

2 Perception and conceptualization of mood variability in folk linguistics

the selection of one mood form over the other by non-linguists in the process of organizing and integrating their everyday experience of the morphological variability in their knowledge of the grammar. It also reveals the structures where the new mood forms are no longer perceived as problematic, a sign that the variability may have given way to a morphological change.

2.1 Different types of “subjunctive errors” and social meanings associated with them

A testimony to the reciprocal influence binding together popular metalinguistic knowledge and language practice is the fact that the subjunctive has been subject to a great deal of comment and reflection from non-linguists and educators alike. This domain of grammar has sparked and continues to spark a lively debate in the public sphere in Italy.

The discussion over the correct and incorrect uses of the subjunctive and the debate over the social meanings associated with language practices resound from private conversations to the media, taking a large place in the newest forms of interactive communication in social media. Among the language structures that have caught the public eye, the subjunctive has reached what Silverstein (2003) calls the third order of indexicality (p. 221): stereotypical language practices associated with subjunctive morphology are consciously embedded into new discourse contexts which presuppose metalinguistic awareness (Agha, 2006; Eckert, 2008; Johnstone, 2016). Using subjunctive morphology

2 Perception and conceptualization of mood variability in folk linguistics

in certain contexts or, conversely, selecting the indicative or the conditional mood to encode semantic functions conveyed by the subjunctive according to the precepts of the prescriptive grammar has become increasingly associated with clichéd identities and behaviors. These language practices are overtly commented or mimicked in a variety of discourse contexts.

Moreover, in the collective imagination the subjunctive has become the most recognizable symbol for the knowledge of the Italian grammar, a sort of indicator of the state of health of the language based on which to measure whether the current generations are able to preserve their linguistic heritage. It figures in the ranking lists of the most common language mistakes produced by native speakers of Italians compiled by various media, like digital newspapers and magazines. Thus, Liberiamo (2016) reports the results of a survey on the Italian language conducted on a sample of 8000 speakers aged between 18 and 65 using Web Opinion Analysis (WOA). According to the survey, 69% of the people interviewed frequently commit the mistake of using an indicative where the standard prescribes the subjunctive. The website concludes that the subjunctive, along with the incorrect use of the apostrophe, is a true “Achilles’ heel” of Italians:

L’uso del congiuntivo (69%) – E qui di sicuro si alzerà un boato. Il congiuntivo, il vero tallone d’Achille di moltissimi studenti e non. Quanti strafalcioni sentiamo ogni giorno anche, e soprattutto, in televisione? “L’importante è che hai superato l’esame”, seppur molto usata questa è una formula

2 Perception and conceptualization of mood variability in folk linguistics

grammaticale scorretta perché in questo caso bisogna usare il congiuntivo: “L’importante è che tu abbia superato l’esame”. [The use of the subjunctive (69%) – This will surely raise a roar. The subjunctive, true Achilles’ heel of many students and people. How many errors do we hear everyday, most of all on TV? “The most important thing is that you passed (ind) the exam”. Although very much used, this grammar form is incorrect because the subjunctive must be used: “The most important thing is that you passed (subj) the exam.”]

The cultural magazine also notes how strongly people generally feel about this type of morphological error and identifies the culprit of the blatant misuse of the mood in the digital communication that has made Italians incapable of writing correct sentences and of exercising critical thinking. The results of the survey have received a great deal of media attention and have been reported also by the major newspapers like *Corriere della Sera* (Redazione Scuola, 2019), *Il Mattino* (2017) and *Il Messaggero* (2017).

Nanopress gets more specific about which forms of the subjunctive paradigm are putatively used incorrectly. The past subjunctive, *congiuntivo imperfetto*, is supposedly misused the most:

Bestia nera della grammatica italiana, il congiuntivo è usato, molto spesso, in maniera errata, soprattutto se si tratta di congiuntivo imperfetto. Uno degli errori grammaticali più comuni riguarda, ad esempio, i verbi *stare* e *dare*: la terza

2 Perception and conceptualization of mood variability in folk linguistics

persona singolare del congiuntivo perfetto è *stesse* e non *stasse*, *desse* e non *dasse*. [Pet peeve of Italian grammar, the subjunctive is used often incorrectly, especially the imperfect subjunctive. One of the most common mistakes concerns the verbs *stare* “stay” and *dare* “give”: the third person of the subjunctive imperfect is *stesse*, not *stasse*, *desse* not *dasse*.] (Padula, 2016)

The newspaper Repubblica online (Mazzocchi, 2016) publicizes a manual written by two scholars and linguists (Patota & Della Valle, 2016) that helps laypeople avoid the most common mistakes and speak correct Italian. The authors suggest that the Italian language is a business card, a tool that shapes the identity of the individual and the image they present to the other. As such, it may open up or forfeit opportunities in the public and work sphere. In the book, the subjunctive figures as one of the most common and most damning mistakes for a person’s reputation, because it betrays lack of knowledge and instruction. Therefore, it must be relearned. The putative mistakes that are especially targeted concern the substitution of past subjunctive forms with conditional forms and the incorrect use of the indicative in lieu of a subjunctive in sentences introduced by subordinating elements such as *malgrado che* “despite that”, *a condizione che* “provided that”.

In social discourse people associate specific identities to the use of subjunctive morphology or the selection of other mood forms in its stead. The social connotations of language practices concerning mood variability are shared within the community and build the substrate for

2 Perception and conceptualization of mood variability in folk linguistics

a plethora of communicative goals in discourse: irony, parody, self-positioning, distancing, building closeness etc. The subjunctive is particularly exploited for ironic or satiric purposes in cultural products like films, music, books which, through open metalinguistic remarks or more often through the direct reenactment of linguistic styles and registers, aim at giving a picture of the Italian society, its customs and its way of life. By tracing this public and fictional narrative around subjunctive errors, it is possible to see how the values and stereotypes associated with the subjunctive have evolved over time possibly alongside usage.

A typical error is the use of the morphological endings of one conjugation with verbs that belong to a different conjugation. This mismatch between verb stem and suffix generates pseudo-subjunctive forms that do not actually exist in Italian, e.g. *nonce* forms **vedino* “see” instead of *vedano* or **vadino* instead of **vadano* for the third person plural of the present subjunctive. The Treccani grammar handbook (Treccani, 2012) notes that most often the subjunctive suffixes of the first conjugations are extended to other verbs, presenting this phenomenon as an instance of morphological overgeneralization of the biggest verb class and a result of analogical simplification. These errors have been greatly exploited in the comedy genre. They occupy a special place in popular imagination at least since Antonio De Curtis, a famous Neapolitan actor and comedian, made them one of the most distinguishable features of the idiolect spoken by Totò, the iconic fictional character he brought to life in numerous popular comedy films,

2 Perception and conceptualization of mood variability in folk linguistics

theatre shows and TV sketches. Some of the most notable pseudo-subjunctives that the actor uses in the Totò film series have been gathered by Rossi (2002) and are reported here:

«ma mi faccino il piacere!» (Totò cerca casa, 1949; Totò e le donne, 1952); «venghi» (detto da o Benti); «passino» (Totò a colori, 1952); «digli che ti dasse» (Totò cerca moglie, 1950); «se ne vadino» (Totò e le donne, 1952); «vadino» (Totò, Vittorio e la dottoressa, 1957; I tartassati, 1959); «faccino» (Totò nella luna, 1958); «mi permettino» (I tartassati, 1959); «Mi facci scendere (detto 4 volte da un personaggio secondario, in Signori si nasce); «venghino» (Totò, Peppino e... la dolce vita, 1961); «se ne vadi» (Il monaco di Monza, 1963). Abbiamo anche un ipercorrettismo: «mi saprebbe dire che significhi Paliatone?» (Totò e i re di Roma, 1952) [“please!” (Totò cerca casa, 1949; Totò e le donne, 1952); “come” (said by Benti); “pass” (Totò a colori, 1952); “tell him to give you” (Totò cerca moglie, 1950); “go away” (Totò e le donne, 1952); “go” (Totò, Vittorio e la dottoressa, 1957; I tartassati, 1959); “make” (Totò nella luna, 1958); “let me”; “Let me come down” (said 4 times by a minor character in Signori si nasce, 1960); come (Totò, Peppino e... la dolce vita, 1961); “go away” (Il monaco di Monza, 1963). There is also a hypercorrection: “could you tell what Paliatone means? (Totò e i re di Roma, 1952)] (pp. 77-79)

2 Perception and conceptualization of mood variability in folk linguistics

The character is an archetype inspired by the Italian *Commedia dell'arte* and his traits, like his language, are a caricature (Anile, 1997, 1998, 2017; Fo, 1995). It is a representative of stereotypical traits that characterize an entire community. The morphological errors index a variety of popular or regional Italian spoken by people coming from the lower class (Rossi, 2002). They serve to depict an uneducated man, coming from poverty who creatively seeks for ways to survive (Bispuri, 1997) and whose efforts to use a high register and impress others with his persona do not quite succeed and usually result in a mixture of high bureaucratic jargon and low vernacular expressions. The distorted macaronic subjunctives in themselves produce a grotesque and comical effect. They are the hallmark of the talk of charlatans, people wanting to pass as cultivated for their own personal advantage. In Totò, however, they can also be understood as an irreverent way of mocking the authority of the upper classes, whose bureaucratic language, a tool for exercising power over the poor and the ignorant, is adopted and at the same time is mocked and made meaningless. It is a criticism of a class who is vacuous and futile like their own language (De Sica, 1957).

The ill-formed subjunctives are well-known among laypeople and go beyond the character of Totò, they embody stereotypical traits which all Italians may have. The phrase “Venghino signori, venghino!” involving a nonce form **venghino* with an exhortative meaning “Come Gentleman, come!” is nowadays a formulaic expression. It evokes the entreaties of a ringmaster inviting the crowd to attend the circus show and is commonly used in all sorts of contexts for satirical purposes,

2 Perception and conceptualization of mood variability in folk linguistics

usually to evoke a fake person, a charlatan, or a fake deal. It is often employed with these connotations in newspaper articles to call out deceit. A few examples are Leone (2018), Sr*, L'Impertinente (2019), *Corriere della Sera* ("Renzi-venditore", 2014), Lauria (2011).

The morphological distortions of the subjunctive forms which characterize Toto's colorful, vernacular-infused language in the 1950s and in the 1960s, a character whose Neapolitan identity was always at the center, have subsequently served to bring to life another iconic character of the Italian comedy. The accountant Ugo Fantozzi is a popular literary and cinematographic figure invented by the writer and comedian Paolo Villaggio. He is the protagonist of a number of best-selling books (Villaggio, 1971, 1974, 1976, 1979, 1983, 1985), stories published in magazines such as *L'Espresso* and *L'Europeo*, highly successful films from the 1970s up until the 1990s and enjoys in Italy a popularity similar to Arthur Conan Doyle's Sherlock Holmes in the English-speaking world.

The pseudo-subjunctives define not only his jargon, but the talk of the entire cohort of people, colleagues and superiors, at the company where the protagonist works in the world of the industrial middle class of Italy's 70s. Fantozzi's universe portrays the mediocre bureaucratic life of the middle-class man in the 70s, of employees and clerks who are subservient to authority, harassed by daily preoccupations and incapable of escaping routine (Pallotta, 2006; Buratto, 2003; Rissa, 2015, pp. 3-39). It is the Italy of the economic boom, of people without a university education who are soon employed in the economic engines

2 Perception and conceptualization of mood variability in folk linguistics

of society and can enjoy relative financial stability. The characters in Fantozzi are also archetypes, *maschere* depicting archetypical behaviors, but, unlike Totò whose regional identity is apparent and who is almost a picaresque figure characterized by a creative eccentricity, they cannot be framed in a specific regional context or be understood as outsiders.

The distorted subjunctives are now the language of the entire middle-class and become the main comic expedient of various film and book scenes. A few notable examples of nonce subjunctives from Fantozzi's cinematographic universe come from two cult scenes, which are reported in the following:

[The inspector at Fantozzi's firm by means of various ridiculous subjunctives politely urges the countess to baptize the firm's ship] "Vadi contessa, vadi!"; "Ri-vadi contessa, ma più centrale!"; "Ri-ri-vadi contessa, ma un po' più a destra!" ["Do it countess, do it!"; "Do it again, but this time aim more to the centre!"; "Do it again countess, but aim more to your right!"] (Villaggio, 1974)

[Exchange between Fantozzi and the accountant Filini, his colleague, while they are about to play a tennis match together] "Facci lei"; "Allora ragioniere che fa? Batti?" – "Ma mi dà del tu?" – "No, no, dicevo, batti lei?" – "Ah! Congiuntivo?" – "Sì!" ["Do it"; "So, what are you waiting for? Are you going to serve? – "Are you addressing me with the informal 'you'?" –

2 Perception and conceptualization of mood variability in folk linguistics

“No, no I meant formal ‘you serve’?” – “Ah! Subjunctive?” –
“That’s right!”] (Villaggio, 1971)

Both scenes are currently available on YouTube, a testimony to how popular and loved by the audience the linguistic irony constructed with the subjunctives actually is. The common trait evoked by using these subjunctives in Totò’s and Fantozzi’s world is the affectation of people’s desire to pass as more sophisticated than they actually are. Since then, public attention towards the subjunctive has not ended. It has rather increased over time. However, language practices that are at the center of the metalinguistic discourse and the stereotypes associated with them have also shifted.

In, *Ovosodo*, the coming-of-age film by Virzì (1997), the subjunctive is presented as a symbol of an unmanly behavior, or sissy talk. It does not belong to the real world, where men conquer life and women through decisive action: “Un congiuntivo in più, un dubbio esistenziale di troppo, ed eri bollato per sempre come finocchio” [An extra subjunctive, one doubt too many and you were forever branded a queen] (reported also in Pietropaolo, 2014, p. 7). The film hints at one of the most common stereotypes evoked in more recent times which associates the subjunctive to gender. Men are pictured as the vandalizers of the standard grammar while women, presented as more educated, verbally gifted and preservers of language still notice or care about the correct use of the subjunctive. It is no longer a certain class, or a certain type who lacks higher instruction and aspires to ascend the social ladder, it is an entire gender, no matter their social background

2 Perception and conceptualization of mood variability in folk linguistics

and the education they received. This connotation was absent in the past.

The stereotype according to which it is mainly male speakers who cannot appropriately use the subjunctive and where such mistakes make the speaker fatally unattractive to the eyes of a female interlocutor becomes the trope of songs, satirical webpages etc. Incredibly well-received is the song “Il Congiuntivo” [Subjunctive] by Lorenzo Baglioni entirely dedicated to the subjunctive mood, presented at the famous national song contest *Festival di Sanremo* in 2018 and reported to be the most popular song of the Festival by Leggo (2018). The song lyrics are given in their entirety below.

Che io sia

Che io fossi

Che io sia stato

Oh oh oh

Oggigiorno chi corteggia incontra sempre più difficoltà

Coi verbi al congiuntivo

Quindi è tempo di riaprire il manuale di grammatica, che è

Che è molto educativo

Gerundio, imperativo

Infinito, indicativo

Molti tempi e molte coniugazioni, ma

2 Perception and conceptualization of mood variability in folk linguistics

Il congiuntivo ha un ruolo distintivo
E si usa per eventi che non sono reali
È relativo a ciò che è soggettivo
E a differenza di altri modi verbali
E adesso che lo sai anche tu
Non lo sbagli più

Nel caso che il periodo sia della tipologia dell'irrealtà (si sa)
Ci vuole il congiuntivo
Tipo "Se tu avessi usato il congiuntivo trapassato
Con lei non sarebbe andata poi male"
Condizionale
Segui la consecutio temporum

Il congiuntivo ha un ruolo distintivo
E si usa per eventi che non sono reali
È relativo a ciò che è soggettivo
E a differenza di altri modi verbali
E adesso ripassiamo un po' di verbi al congiuntivo

Che io sia (presente)
Che io fossi (imperfetto)
Che io sia stato (passato)
Che fossi stato (trapassato)
Che io abbia (presente)
Che io avessi (imperfetto)

2 Perception and conceptualization of mood variability in folk linguistics

Che abbia avuto (passato)

Che avessi avuto (trapassato)

Che io sarei

Il congiuntivo, come ti dicevo

Si usa in questo tipo di costrutto sintattico

Dubitativo, quasi riflessivo

Descritto dal seguente esempio didattico

E adesso che lo sai anche tu

Non lo sbagli più

[That I be

That I were

That I have been

Oh oh oh

Nowadays suitors face more and more difficulties

With subjunctive verb forms

Therefore it is time to open again the grammar textbook

which is

Very instructive

Gerund, imperative

Infinitive, indicative

Many tenses and many conjugations, but

The subjunctive

Has a distinctive role

2 Perception and conceptualization of mood variability in folk linguistics

And it is used for events that aren't real

It's related

To what's subjective

Unlike other verb moods

And now that you know it too

You'll no longer get it wrong

If the sentence encodes an unreal situation (it's common knowledge)

The subjunctive must be used

Like that "If you had used the pluperfect subjunctive

It wouldn't have ended so miserably with her

Conditional

Follow the correct sequence of tenses

The subjunctive

Has a distinctive role

And it is used for events that aren't real

It's related

To what's subjective

Unlike other verb moods

And now let's review some subjunctive forms

That I be (present simple)

That I were (past simple)

That I have been (present perfect)

2 Perception and conceptualization of mood variability in folk linguistics

That I had been (past perfect)

That I have (present simple)

That I had (past simple)

That I have had (present perfect)

That I had had (past perfect)

That I would be

The subjunctive, as I was saying

Is used with this syntactic construction

Doubtful, almost thoughtful

Described by following example

And now that you know it too

You'll no longer get it wrong]

(Baglioni, 2018, *Il Congiuntivo* [Song] on *Bella, Prof!*
[Album], Sony Music)

The song video which boasts more than 7 million views on YouTube at the time of writing (Baglioni, 2018) narrates the story of a young man who falls in love with a woman and dedicates her a card with the message in (9).

(9) Se io starei con te, sarei felice.

If I be.1.SG.PRS.COND with you, be.1.SG.PRS.COND happy.

“If I were with you, I’d be happy”

2 Perception and conceptualization of mood variability in folk linguistics

The sentence exhibits a typical morphological error in the conditional construction: the conditional mood is used instead of the prescribed subjunctive in the protasis. The error is a result of the distinctive tendency in speech to level out morphological asymmetries and use the same mood and tense in both clauses of a hypothetical construction. The form of one of the two verbs usually influences the other (Harris, 1986; Sankoff & Laberge, 1978; Leblanc, 2009). The message written on a heart-shaped card causes the young woman to run away from her suitor after she squints at the message incredulous and horrified at the mood form used. “If you had used a subjunctive, it wouldn’t have gone so miserably with that girl”, sings the author using a sequence of verbs with an impeccable mood combination and a correct *consecutio temporum* in the conditional construction: “Se tu avessi usato (PPRF SUBJ) il congiuntivo trapassato / Con lei non sarebbe andata (PST COND) poi male”.

In the song video, the young man who is violently rejected because of his morphosyntactic blunder is not defined by any characteristic trait that might reveal lack of education or a particular sociocultural background. He is a typical well-groomed university-educated millennial. Indeed, the song text remarks that nowadays people who are looking for their soul mate encounter more and more difficulty with the subjunctive: “Oggigiorno chi corteggia incontra sempre più difficoltà /Coi verbi al congiuntivo”.

The explicitly stated purpose of the satiric ballad is to teach when and how to use the subjunctive and refresh the subjunctive forms for

2 Perception and conceptualization of mood variability in folk linguistics

the listener. The song demonstrates the continued relevance of the subjunctive in public perception and how strongly laypeople feel about it today. The attention, however, as also Della Valle and Patota (2011) remark, shifts from the morphology to the syntax of the subjunctive. It is no longer the incongruous stem-affix combinations that are mocked but the mood alternation in precise syntactic contexts.

On YouTube, the song has received thousands and thousands of positive comments with users invoking that the song be listened to in schools for educational purposes or be made the national anthem (Baglioni, 2018). Countless users testify that it was indeed their teacher of Italian who made them listen to the song in school. The hymn to the subjunctive is being used for educational purposes not only in the context of the domestic national education but also by non-native speakers who are learning Italian around the world (Baglioni, 2018).

The public particularly relates to the stereotype associating mood to gender. One female user comments that she dated a man who once used another mood form instead of the subjunctive. She sent him the video and they became a couple: “Un ragazzo col quale mi sentivo un giorno aveva sbagliato un congiuntivo, gli ho inviato questo video, ora è il mio ragazzo. Viva la *consecutio temporum* e viva sta canzone” (Baglioni, 2018). Another female user remarks that she is not demanding in terms of the qualities that her own ideal boyfriend should possess; the important thing is that he knows how to use the subjunctive (Baglioni, 2018). A male user wonders whether he can finally seduce the woman of his dreams showing up at her door with signs parading correct

2 Perception and conceptualization of mood variability in folk linguistics

subjunctive forms (Baglioni, 2018). The public has thus developed an understanding of the phenomenon and a rhetoric to present it which is conceptualized in the principle first developed by Labov (2001) on the social dimensions of language variation. The principle is based on large evidence coming from different types of communities (Fischer, 1958; Wolfram & Fasold, 1974; P. Trudgill & S. Trudgill, 1974; Labov, 1966; Milroy & Milroy, 1978 among others) and states that: “For stable sociolinguistic variables, women show a lower rate of stigmatized variants and a higher rate of prestige variants than men” (p. 266). Labov (2001) concludes in his famous gender paradox that “[w]omen conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not” (p. 293). Women, according to the evidence coming from different languages, are more sensitive to matters of prestige in language and tend to avoid highly stigmatized mood forms, a behavior that, in the case of the subjunctive, has not escaped the public.

The attention that the subjunctive is enjoying in the social life of Italians is especially tangible in social media. In particular, after Baglioni’s song, twitter users started to embrace the cause of the subjunctive as a social cause on twitter employing the following hashtags: *#JeSuisCongiuntivo* (echoing the famous slogan advocating for freedom of speech *#JeSuisCharlie*), *#CheIoSia*, *#CheIoFossi*, *#WilCongiuntivo*. Even long before the more recent call to arms, entire webpages had been dedicated to the issue of mood variability and to debating the errors concerning the subjunctive, a testament to the

2 Perception and conceptualization of mood variability in folk linguistics

importance attributed to language as a common heritage and asset, and compelling evidence of how strongly people feel about language use and grammar.

A number of Facebook webpages are devoted to reporting linguistic and social phenomena associated with the subjunctive (see “Correggere mentalmente l'interlocutore con congiuntivi e condizionali”, 2010; “Il congiuntivo non è una malattia degli occhi”, 2010 among others). These websites are very popular, with a support ranging from tens of thousands to hundreds of thousands of active users. The Facebook webpage “Correggere mentalmente l'interlocutore con congiuntivi e condizionali” (2010) [Mentally correcting the subjunctives and conditionals of your interlocutor] mocks Italians who cannot properly use the subjunctive. The stereotype invoking a gender dichotomy in the morphological competence is given a central relevance in the page. The cover image portrays men as fatally unattractive because they speak a “bad language” or a “bad grammar”: “finally find real love, unable to use subjunctive” reads the caption of the image and frames the picture of a disappointed heart-broken woman and an ashamed and remorseful man. The message that the website successfully conveys is no matter how good looking you are, if you cannot speak a correct Italian, nobody will like you. The purport behind the joke is apparent and shared among the community at large because it draws on very well-known grievances. As already highlighted by the comment of a YouTube user to the song “Il Congiuntivo”, in social interactions Italians often complain about having met a potential partner and being disappointed

2 Perception and conceptualization of mood variability in folk linguistics

by the poor Italian he or she speaks, perceived as an indicator of ignorance and coarseness. The cartoon “Azzecca un congiuntivo!” [Get a subjunctive right!] (“Il congiuntivo non è una malattia degli occhi”, 2015) also depicts the stereotypical gender divide in the knowledge and correct use of the subjunctive. A woman in tears implores her lover to “get a subjunctive right!”.

The cartoon titled “Distanziamento sociale” [Social distancing] (Caon, 2020) and posted on various social media has a similar subtext relating language behaviors to desirability. It refers to the social distancing measure (of at least 1.5 meters between individuals) required by a number of national governments throughout the world for preventing the spread of the viral disease during the Covid19 pandemic in 2020. In parallel, a distance of 10 meters from people who incorrectly use the conditional in lieu of a subjunctive in the if-clause is required according to the cartoon. The joke has two interpretations: on the one hand people who make this grammatical mistake are undesirable and unattractive, one cannot be near them. On the other hand, the spread of the conditional mood in subjunctive clauses is so viral that one should keep a guarded distance in order to avoid being infected by the corrupting linguistic disease.

The Facebook page “Il congiuntivo non è una malattia degli occhi” (2010) [Subjunctive is not an eye disease] reports morphological mistakes related to the subjunctive. With 224 thousand fans at the time of writing, the page laments that Italians have completely erased the subjunctive from their linguistic knowledge: “Subjunctive is not an eye

2 Perception and conceptualization of mood variability in folk linguistics

disease” states the cover title, a word pun bringing together the terms subjunctive and conjunctivitis, which in Italian are nearly homophonous: *congiuntivo* and *conjuntivite*. The title mocks the general confusion and ignorance of the population who is not even able to tell the difference between the two concepts anymore. These types of pages or groups abound in social media. They usually have a satirical purpose but they also engage in a call to arms to mobilize the public for the issue of the subjunctive. In the group “Lottiamo contro la scomparsa del congiuntivo” (2007) [Let us fight against the disappearance of the subjunctive], the mood is treated like an endangered species that must be protected by spreading linguistic awareness.

The title of the satirical Facebook page “Correggere mentalmente l'interlocutore con congiuntivi e condizionali” (2010) makes an interesting reference to the linguistic practices of the community stemming from language normativism. People correct their interlocutors and get corrected in turn all the time for an improper use of the verb moods during conversations. The natural flow of the conversation is thus sometimes interrupted in order to allow for meta-linguistic comments, to reproach people or to correct them. This behavior suggests that the subjunctive is cognitively highly salient, even though non-standard morphological forms are commonly produced. Valuable evidence of this peculiar phenomenon which can be witnessed in private conversations are the debates and discussions available in online chat forums. A user poses a particularly relevant question on Yahoo Answers: “Why does everybody make mistakes

2 Perception and conceptualization of mood variability in folk linguistics

when it comes to the subjunctive? Do you correct them?”. The excerpt is reported in the following.

Perché tutti sbagliano i congiuntivi? E voi li correggete? Ieri che imbarazzo [...], un dottore che misurava l'udito stava parlando con mio padre e a un certo punto ha detto “se avrei...”. Mio padre gli ha detto: “scusi, un dottore dovrebbe sapere che si dice *se avessi*, io lo dico solo per poterla aiutare”. Io sono diventata viola all'inizio, poi mi sono accorta che la mia pelle si evolveva in uno strano verde shrek, poi in un blu puffo. Voi, se qualcuno sbaglia un congiuntivo, lo correggete con tutta questa disinvoltura o vi fate i cavoli vostri? E poi, perché tutti li sbagliano? Voi li sbagliate? [Why does everybody make mistakes with subjunctives? Do you correct them? Yesterday I was so embarrassed. A doctor who was performing a hearing test was talking to my dad and suddenly said *se avrei* (COND) “if I would have”. My dad told him: “Pardon me, a doctor should know that the correct form is *se avessi* (SUBJ) “if I had”. I’m telling you this only to help you.” I turned violet, then my skin turned green, then blue. When somebody makes mistakes with subjunctives, do you correct them so casually or do you mind your own business? Why does everybody make this type of errors? Do you make them?] (Yahoo, 2010)

2 Perception and conceptualization of mood variability in folk linguistics

The numerous answers of users vary from the admission that they indeed correct their interlocutors all the time, no matter whether they are close friends or complete strangers, on the grounds that the Italian language must be protected, to the frustration at hearing those very same mistakes and not commenting on them due to politeness, or in order to avoid coming across as snobbish or hypercritical.

Oddio, io sono odiato da tutti per questo motivo. Ho l'istinto irrefrenabile di correggere ogni congiuntivo o condizionale che sento sbagliato. [...] come si può dire *se avrei*? Mi viene un brivido sulla schiena come se stessero graffiando una lavagna! [...] suona malissimo. [Oh, my, everybody despises me for this. I have this irresistible urge to correct every subjunctive and conditional error. How can anybody say *se avrei* (COND)? I get chills down my spine as if somebody were grazing a blackboard. It sounds terrible.] (Yahoo, 2010)

In another comment, a person wonders at the fact that even his teacher of Italian in junior high school frequently uttered incorrect morphological forms, precisely the person who tested the pupils' grammar knowledge

Dipende, di solito correggo se si tratta di amici o parenti, ma nel caso del dottore forse non avrei detto niente. Io avevo la prof di italiano delle medie che sbagliava i congiuntivi e ogni volta a me ad una mia compagna si attorcigliava lo stomaco!

2 Perception and conceptualization of mood variability in folk linguistics

Come fa una prof di italiano a sbagliare i congiuntivi? Secondo me la gente oggi continua a sbagliare i congiuntivi per colpa della television. Il peggio sono i film. Riescono sempre a sbagliare i congiuntivi nonostante tutte le prove che fanno! [...] Anche i politici oggi fanno errori di grammatica. Povero italiano! [It depends, I usually correct friends or relatives, but I would not have corrected a doctor. I had a teacher of Italian in junior high school who always made mistakes with subjunctives and hearing them made our stomach twist. How can a teacher of Italian make such errors? It believe that people nowadays make them due to the TV. Films are the worst. They are full of these errors although they record numerous takes! [...] Also politicians make so many grammatical errors. Poor Italian language!] (Yahoo, 2010)

Notice again one female user commenting that she had the habit of correcting her ex-boyfriend's subjunctives all the time, an observation that reinforces the gender dichotomy:

Conoscere e sapere usare bene i congiuntivi è una discriminante tra chi usa un buon italiano e chi no [...]. Riguardo alla correzione, beh dipende dalla persona in questione, io mi diverto a correggere il mio ex e tra noi ormai era un giochetto; correggo i miei amici [...]. [Knowing well subjunctive forms and using them properly is critical to speaking a good Italian. With regard to the issue of correcting

2 Perception and conceptualization of mood variability in folk linguistics

other people, it depends on the person, I used to correct my ex-boyfriend and it was our thing; I correct my friends.] (Yahoo, 2010)

The debate topic was prompted by a particular real-life incident in which the user and her father talking to a doctor in the context of a medical examination heard him produce the conditional mood form of the verb *avere* instead of the subjunctive in the protasis of a conditional sentence. The utterance is reported in (10).

- (10) Se io avrei
If I have.1SG.PRS.COND
“If I had”

One user participating in the debate points out that the expression in (10) gives him the shivers. The feeling is by no means an exaggeration of a single individual. The satirical webpage “Il congiuntivo non è una malattia degli occhi” (2014) posted a meme on facebook which depicts the same locution with verb *avere* causing the very same skin reaction.

It becomes apparent that the substitution of the subjunctive with conditional forms in this particular syntactic construction causes a strong rejection in general. People relate to the feeling of cringing evoked in the picture. It is the most discussed and incriminating mistake, one that certainly makes headlines when public figures commit it. The famous and highly respected linguistic academy for the Italian language *Accademia della Crusca* (2014) posted on its official

2 Perception and conceptualization of mood variability in folk linguistics

facebook page an authentic photo of a real-life incident: a banner hanging on a bridge with a declaration of love written on top of it. The love message is very similar to the one presented in the fiction of the song “Il Congiuntivo” (Baglioni, 2018) and presents the same critical error (11).

- (11) E se ti amerei per sempre?
And if you.DAT love.1.SG.PRS.COND for ever?
“And if I loved you forever?”

In the caption to the photo, the linguistic society, quoting Biagi (1987, p. 770), writes that once the uses of the subjunctive have been properly learned, the speaker is free to dispense with it and choose other mood forms. Knowing the grammar, however, means being able to choose the speech register that is appropriate to the communicative context according to the academy. The message reflects the tension between normativism and innovation. In the picture, the big gesture of a very public love declaration is ruined by the bad grammar. The non-prescriptivist stance taken by the academy is somewhat weakened by the stigmatization of the use of the conditional mood form in such a context. The improper use of the conditional makes an important moment become ludicrous and the academy highlights the ridicule. Again, a clear stereotype is consciously enacted: the correct use of the subjunctive mood is seen in its elegance in relation to courtship, to how a suitor presents his or her identity in order to charm someone else. It

2 Perception and conceptualization of mood variability in folk linguistics

becomes the symbol of polished language needed for others to notice our fine qualities and approve of us. Indeed, in the numerous comments on the picture, nearly two hundred, some users associated aesthetic values with the subjunctive in addition to moral ones stating among others that the subjunctive is beautiful and refined (Accademia della Crusca, 2014).

Beyond the overarching cliché according to which men in general cannot navigate the grammar and the functions of the subjunctive, one group is particularly condemned for the use of norm-violating mood forms. This is the political class whose ignorance of the subjunctive has become something of an archetype. In the comments on the YouTube video of the subjunctive song some users suggest that politicians like Di Maio, the Minister of Foreign Affairs at the time of writing, should listen to the song, or are among those users that disliked the video (Baglioni, 2018).

Senator Gianluca Castaldi receives public flogging for an incorrect mood form while delivering his speech in the Italian Senate (“Il Senato ha corretto un congiuntivo, in coro”, 2014). The reputation-destroying blunder is the usual one: he does away with an imperfect subjunctive form in favor of the conditional in the protasis of a conditional construction which in this case has the illocutionary force of a request. In (12) I report his utterance.

- (12) Se potrebbe cortesemente controllare.
If can.3.SG.PRS.COND kindly check.INF

2 Perception and conceptualization of mood variability in folk linguistics

“If you may kindly check.”

The interesting fact is that the politician gets immediately corrected by the entire Senate, a chorus of voices raised by the other members who all at once chant the right subjunctive form *potesse* (IMPF SUBJ). The embarrassing incident was reported by all major media outlets along with the footage of the choral correction. The reaction of the Senate, pointing out the error out loud instead of overlooking it, echoes the reports found on the web and discussed above according to which it is common practice to correct one’s interlocutor in such cases. The error does not go unnoticed and most importantly it does not go uncommented both in the private and in the public sphere.

Della Valle and Patota (2011) remark that there is a high sensitivity for the linguistic correctness of the people representing the institutions: “La sensibilità condivisa per la correttezza linguistica della comunicazione istituzionale è molto alta” (p. 16). This echoes the position of the Accademia della Crusca. Although the academy nowadays adopts an essentially non-prescriptivist attitude opposed to linguistic purism, it points out that the norm-violating mood forms reveal a lack of socio-pragmatic competence, a failure to adopt a register that matches the discourse context. Yet politicians, even when they are highly vigilant, while delivering a scripted speech for instance, commit these morphological slips. The case of the Senator is by no means isolated. Years earlier, during an interview with a national TV news broadcaster, the Minister of Education at the time, talking about

2 Perception and conceptualization of mood variability in folk linguistics

the proposal to adopt a bill abrogating the legal distinction between different types of schools of higher education, expressed his wish to start the debate with the utterance reported in (13).

- (13) Vorrei che ne parliamo.
wish.1.SG.PRS.COND that PRO.OBJ talk.PL.PRS.SUBJ/IND
“I wish we talked about it.”

The alleged lack of concordance between the mood and tense forms of the main and embedded clause was blasted in the media. A few days after the incident, *Il Corriere della Sera*, the newspaper with the highest circulation in Italy, published a brief comment titled *Ora abolirà il congiuntivo* [He (The minister D’Onofrio) is going to abrogate the subjunctive next] (Palumbo, 2016). This was only one of numerous harsh reactions. That precisely the minister of education proposing a reform of the school system should make such an ignorant grammatical blunder was according to the media rhetoric and the public perception unforgivable. It manifested the decay and ignorance of the political class, a group made of charlatans (recall the effect of the distorted subjunctives in Totò and Fantozzi) who are in no position to make the most important decisions affecting the country.

The minister confronted with the wave of criticism and ridicule felt compelled to explicitly defend his morphological choice noting that he had indeed used a subjunctive. Only, the affixes encoding the first person plural of the subjunctive and indicative present are identical:

2 Perception and conceptualization of mood variability in folk linguistics

“Non è colpa mia se la prima persona plurale tanto dell’indicativo che del congiuntivo presente sono uguali: parliamo” [It is not my fault if the first person plural of the subjunctive and indicative present are identical], he told the media (“D’Onofrio: ‘Esperti, ditemi se è sbagliato il mio congiuntivo’”, 1994). The defense proved even more damning because the grammatical error according to the critics lied in the *consecutio temporum*, the concordance of tenses between the matrix clause and the embedded clause. The criticism was based on the perception that the conditional form of the verb *volere* in the main clause exclusively requires the imperfect tense of the subjunctive in the embedded complement clause. The present subjunctive is perceived ungrammatical. Sentence (14) exhibits the correct concordance according to the critics.

- (14) Vorrei che ne parlassimo
wish.1.SG.PRS.COND that PRO.OBJ talk.1.PL.IMPV.SUBJ
“I wish we talked about it.”

This incident is particularly interesting because, although the utterance of the minister provoked a very strong and unanimous rejection, the possibility to have the present subjunctive in a complement clause governed by a conditional present form is actually reported in the main grammars of Italian, also with governor verb *volere*.

2 Perception and conceptualization of mood variability in folk linguistics

The rules of the *consecutio temporum* for the temporal agreement between the verbs of the main and embedded clause state that the subjunctive present governed by a present conditional verb encodes a relationship of simultaneity between the events. In (15) are two example sentences taken from Treccani as illustration.

(15)

a. Immaginerei che egli faccia bene.

think.PRS.COND that he do.PRS.SUBJ good

“I think he is right.”(Simultaneity in the present)

b. Immaginavo che facesse bene.

think.1SG.IMPF.IND that do.3SG.IMPF.SUBJ good

“I thought that he was right.” (Simultaneity in the past)

This combination of mood and tense forms is also not unseen in the literature and can be found precisely with the matrix verb *volere* in a number of authors. A few examples are reported in the following with the relevant verb forms highlighted in bold:

Vorrei (PRS COND) che questa donna **sia** (PRS SUBJ) avvisata di non seccarmi! Ho portati con me alcuni libri e vorrei essere lasciato in pace. [I would like for this woman not to bother me! I have brought with me some books and I would like to be left alone.] (Svevo, *La Coscienza di Zeno*)

2 Perception and conceptualization of mood variability in folk linguistics

Come si vede, tutte le mie preoccupazioni riguardano l'immagine di strenuo contestatore che **vorrei** (PRS COND) che Silvia **abbia** (PRS SUBJ) di me. [As you can see, I care about my image as a tireless protester that I would like Silvia to have of me.] (Moravia, *L'uomo che guarda*)

A ogni modo io non vorrei imbrogli, e **vorrei** (PRS COND) che tutto **sia** (PRS SUBJ) pronto per l'8... [Anyway, I would not like any cheating, I would like for everything to be ready for the 8th...] (Carducci, *Lettere*)

The website *Cruscate* dedicated to discussions on the Italian language reports more of such uses. In the grammar of Italian by Serianni (1989), the author specifies however that, although the subjunctive present may be found in complement clauses and other types of embedded sentences governed by a conditional present, the subjunctive past tense is used more often with verbs that encode the semantics of volition: “Se nella reggente figura il condizionale presente di un verbo indicante volontà, desiderio, opportunità (come *volere, desiderare, pretendere, esser conveniente* e simili;), la dipendente si costruisce col congiuntivo imperfetto più spesso che col congiuntivo presente” (Serianni, 1989, p. 562). Such use is more frequent but not categorical according to Serianni. The linguistic academy Accademia della Crusca, however, asked to weigh in on the rules of the *consecution temporum* concerning the subjunctive in the embedded predication, takes the passage by Serianni to make a brief statement that the subjunctive past is more

2 Perception and conceptualization of mood variability in folk linguistics

appropriate. Indeed, most grammar textbooks do not report the use of the present subjunctive any longer and state that only the past tense in embedded clauses governed by a conditional form is grammatical. The fact that only the past tense may be perceived as grammatical nowadays and indeed is the only form prescribed by most grammars adopted in schools may be understood as part of a general trend according to which the subjunctive progressively loses the ability to encode temporal information (Leblanc, 2009, p. 1004). The particular shift in Italian which sees the subjunctive present as no longer grammatical under a conditional mood form must have taken place very recently (see the use in Moravia). The obligatorification of one tense of the paradigm, an example of restriction of the intraparadigmatic variability in a particular syntactic context, is taken as evidence for the larger process of grammaticalization and fossilization by Lehmann (2015, p. 124). According to him, in the grammaticalization cline, “the freedom of the language user with regard to the paradigm as a whole” is compromised (p. 124).

In the case of the minister D’Onofrio, an inspector at the ministry for education, Katia Petruzzi, was the only person to come to his defense thereby opening up the debate on the functional contribution of the subjunctive among non-linguists. She argued that the minister’s morphological choice was legitimate because the subjunctive present suggests immediacy and feasibility, while the subjunctive past gives a connotation of eventuality und irreality to the event. The minister, with a polite form expressed by the conditional *vorrei*, that modulates and

2 Perception and conceptualization of mood variability in folk linguistics

attenuates the illocutionary force of the verb *volere* “to want”, expresses nonetheless his wish that the bill be discussed now and not in a future or ideal world removed from the present (see (Palumbo, 2016). Petruzzi’s argument makes a distinction between the subjunctive present and past that hinges both on the modal and the temporal interpretation. According to her argument, the subjunctive can still be understood and used as a temporal deictic and there is a relationship between the temporal value and the modal one: something that can be done now is also something that is highly possible to do. This intervention was perceived by the media and the general public as a feeble attempt to salvage the minister’s reputation and no source reporting the incident up to these days has dived into the use of the subjunctive present governed by a conditional present form, not even Della Valle and Patota (2011). In most cases, it is no longer accepted.

The list of articles and comments from media sources, pundits and writers in general on the political talk that misuses the subjunctive is endless. Political figures of great importance in the public sphere, like Di Maio, Conte, or Di Pietro are routinely accused of making such morphological blunders in speech and writing (Francese, 2018; F.Q., 2018; Agenzia Vista, 2018; La Repubblica, 2017; Fazio, 2018; Antonelli, 2019). Prime minister, Giuseppe Conte, has been called Fantozzi by the press (Il Giornale, 2018) for uttering a series of pseudo-subjunctives that recall the ludicrous language of Villaggio’s famous character (Villaggio, 1971, 1974) discussed above. Some newspaper and online articles like “La lotta con il congiuntivo: Di Maio soltanto

2 Perception and conceptualization of mood variability in folk linguistics

l'ultimo della catena" (Palumbo, 2016), "L'agonia della lingua italiana in politica" (Ferrara, 2017) and others (Terranova, 2017) review all the most notable cases, some of which have been reported here, along with the reactions to them. They mention among others one instance where a politician and former premier, Romano Prodi, took to the press, specifically in a piece published by the *Corriere della Sera*, to lament that the opposing party, *Il Polo per le Libertà*, had stolen the political program of his own group, the only difference were the subjunctive errors in the copy: "Ci sono intere frasi identiche, ma con qualche congiuntivo sbagliato, con un italiano peggiore" (Palumbo, 2016). The linguist Luciano Satta is reported to be the author of a famous pronouncement: "Il potere logora i congiuntivi di chi lo detiene" [Power consumes the subjunctives of those who have it] (in Antonelli, 2019).

Yet another satirical song from the website and satirical blog *La Sora Cesira* (2014) has been dedicated specifically to the tormented relationship between the subjunctive and the political class. The song video (La Sora Cesira, 2018) presents a compilation of video clips in which politicians make morphological mistakes involving the subjunctive while delivering a speech in parliament or during talk shows and interviews. Most of the errors filmed are distorted forms of the subjunctive present, where the affixes of the 1st conjugation are used together with verbs of the 2nd or other conjugations: *scrivino* (instead of *scrivano* "they write"), *arricchischino* (instead of *arricchiscano* "they enrich"), *facci* (instead of *faccia* "you do"), but also distortions of

2 Perception and conceptualization of mood variability in folk linguistics

subjunctive past forms. In this case, the inverse process takes place: the affixes of the 2nd conjugation are used together with verbs of the 1st conjugation and create the following nonce forms: *assocessimo* (instead of *associassimo* “we associate”), *troveressimo* (instead of *trovassimo* “we find”). Previously, it was discussed how these types of errors have long been exploited in fictional contexts for ironic purposes. However, although they are highly salient in metalinguistic discourse, stereotypical even, they still occur in non-fictional talk. Additional errors recorded in the talk of politicians are the substitution of the subjunctive past with the conditional mood in hypothetical sentences and the use of the subjunctive when other mood forms are required, for instance in embedded complement clauses governed by the matrix verb *dire*.

The picture that emerges from the song clip and the articles is one of strong confusion and clumsiness in the political speech. The public figures have a hard time uttering the correct form. Sometimes they attempt to produce it multiple times. Most importantly, the pervasive morphological variation is perceived as an index of decadence of the political class, of the ignorance and incompetence of the institutions in charge of guiding the country. The political class in turn reflects the general state of decay of modern times from a past of classical grandeur that is usually attributed to both Italy and the Italian language, referred to in these contexts with a common epithet as the language of Dante (the song text also uses this particular epithet to frame the ironic narration). Notice that all the political figures whose morphological

2 Perception and conceptualization of mood variability in folk linguistics

errors are reported are men. Not one woman is present among the politicians accused of vandalizing the grammar of Italian.

The writer and journalist Rizzo (2016) reviewing contemporary Italian politics makes the main argument that the political class is in decline. Beyond the corruption, the personal interests and the lack of morals, he argues that the lack of education as a result of a society that has given up on the school system as its fundamental pillar is the reason why the ruling class is incapable of pursuing public interest as its North Star. The book is dedicated to gathering evidence for this hypothesis, and, among cases of corruption and poor management, there is a section dedicated to the ignorance of the Italian language. The title reads “Maledetto congiuntivo” [Freaking subjunctive!]. The subjunctive is again in the eye of the storm; it is taken as evidence that the power is held by mediocre individuals. In the public rhetoric, the morphological variation and simplification reflect the social, cultural and economic crisis. In this light, there is a continuum between the morphological slip occurring in day-to-day interactions, which betrays the ignorance of the average man and voter, and the error-ridden talk of the highest institutional ranks, seen as the product of the uneducated masses, those that elected their representatives. To this regard see Gramellini’s (2007) article on *La Stampa* claiming that an incorrect mood form substituting the subjunctive was used on purpose in a Radio commercial commissioned by the Ministry of Labor to appeal to the ignorance of the general public. Similar arguments are made in a book entirely dedicated to the language of the political class by Antonelli (2019).

2 Perception and conceptualization of mood variability in folk linguistics

Yet, the public feels strongly about the subjunctive. Linguistic normativism concerning the use of the subjunctive mood exercises a strong pull in the collective conscience. For this reason, as it has been shown in this chapter, a variety of social meanings are attached to those phenomena of morphological variation and simplification that trigger the radar of grammatical correctness of non-linguists. They are then exploited in social discourse for a variety of purposes: satire, public shaming, building arguments about society and culture in general, positioning oneself and negotiating one's identity.

The metalinguistic discourse in social media and the more traditional media makes sure that the subjunctive is mentioned cyclically. Della Valle and Patota (2011) calculated an average of ten mentions each year for each major media outlet. For this reason, they coined the expression *questione del congiuntivo* “the subjunctive issue”. The locution aims to indicate that the subjunctive is a real issue for Italians, one that has been around for a long time and that cannot be easily resolved.

2.2 Conclusion

Based on the folk perception of mood variability reported above, a number of important observations on morphological variation and simplification affecting the use of the subjunctive can be made. First, there is societal pressure to conform to the prescribed use of the subjunctive at least in certain syntactic contexts. The norm-violating practices that receive the greatest amount of attention are:

2 Perception and conceptualization of mood variability in folk linguistics

1. The substitution of the subjunctive imperfect with the conditional present mood in the protasis of conditional sentences of the second and third type called *conditionals of possibility* and *conditional irrealis* and usually introduced by conjunction *se* “if”

In (16) I report, as an example, Senator Castaldi’s utterance already given in (12).

(16) Se potrebbe cortesemente controllare.
If can.3.SG.PRS.COND kindly check.INF
“If you may kindly check.”

2. The production of illicit morphological forms resulting from the combination of morphemes and verb stems that do not belong to the same conjugation class, as in (17) (recall *facci* in lieu of *faccia* for the singular forms of present subjunctive of *fare* “do”).

(17) Può darsi che gliel’**abbino** data
May be.INF that him-have.#3PL.PRS.SUBJ give.PP

perché qui non c’è nulla.
because here NEG ADV-be.3SG.IND nothing
“It is possible that they gave it to her because here there’s
nothing.” (C-ORAL-ROM, ifamcv01)

2 Perception and conceptualization of mood variability in folk linguistics

The use of the conditional in lieu of a subjunctive imperfect form in if-clauses as in (16) is still marked, at least in some social groups. So is the use of the imperfect indicative in this syntactic context, even though it is perhaps less criticized than the conditional. (18) is an example taken from the C-ORAL-ROM corpus.

- (18) Io non avevo nessun problema
I NEG have.IMPF.IND no problem

se venivate anche voi.
if come.IMPF.IND also you

“I wouldn’t have had any problem if you had come too.”
(ifamcv08)

Following Bybee & Hopper (2001) and Bybee (2010), it could be hypothesized that the subjunctive is part of a linguistic chunk in this instance. In such cases, other frequently co-occurring elements of the clause, such as the conjunction *se*, would trigger its selection. However, it is possible to notice a resistance towards the substitution of the imperfect subjunctive in general and especially if the replacement is a conditional form. Recall how in (13) the subjunctive present occurring in a complement clause governed by the conditional form of volitive verb *volere* was strongly rejected. The subjunctive imperfect was claimed to be the only grammatical option. This sentence is reported again in (19).

2 Perception and conceptualization of mood variability in folk linguistics

- (19) Vorrei che ne parliamo
wish.1.SG.PRS.COND that PRO-OBJ talk.1.PL.PRS.SUBJ/IND
“I wish we talked about it.”

The morphological variability encountered in speech, however, does not only affect the contexts listed above. The substitution of the subjunctive in complement clauses, usually with the indicative, is still one of the most common phenomena, as evidence from corpus studies suggest (Poplack, 1990; Poplack et al., 2018 and my study in chapter 5). Examples of the selection of the indicative present in lieu of the subjunctive present in complement clauses are given in (20) and (21).

- (20) Basta che uno fa
Suffice.3.SG.PRS.IND that one do.3.SG.PRS.IND

una stupidaggine.
a stupidity
“It suffices that one commits a dumb thing.” (ilfamnn02, C-ORAL-ROM)

- (21) Spero che in questi otto anni
Hope.1.SG.PRS.IND that in these eight years

gli occhi del mondo possono

2 Perception and conceptualization of mood variability in folk linguistics

the eyes of-DET world can.3.PL.PRS.IND

contribuire a migliorare la situazione [...]

contribute.INF to improve.INF the situation [...]

“I hope that in these eight years the eyes of the world can contribute to improve the situation [...]”. (imedts07. C-ORAL-ROM)

Overall, this type of variation is far less commented, although some mention of such practices can still be encountered. Again, somewhat more marked is the substitution of the subjunctive imperfect, also in complement clauses, as in (22).

(22) Volevo che ci raccontavano questa cosa.

Want.IMP.F.IND that PRO tell.IMP.F.IND this thing

“I wanted them to tell us that thing.” (ifamcv17, C-ORAL-ROM)

These facts suggest that the imperfect is more stable than the present tense or that at least it is more salient in folk perception. Concerning the new, “non-standard” morphological variants, they also indicate that, overall, the indicative is perceived as more acceptable than the conditional, at least in certain contexts.

It is interesting to note that the diachronic development of the subjunctive system in Italian diverges from the evolution of the French *subjunctif*. As a matter of fact, speech practices signal quite the opposite

2 Perception and conceptualization of mood variability in folk linguistics

trend in contemporary French where the subjunctive present is still more productive in speech. It accounts for all the occurrences in the corpus surveyed by Poplack et al. (2013). On the other hand, the subjunctive imperfect has now nearly died out in spoken French. In fact, in standard French, the conditional construction, which receives great attention in Italy, does not exhibit the same morphological structure. The if-clause (protasis) of the hypothetical construction requires no subjunctive. Instead, the imperfect indicative is prescribed. The conditional appears in the apodosis. This morphosyntactic combination can be found as a non-standard and stigmatized variant in Italian. Indeed, in informal speech, the imperfect indicative occurs frequently in the if-clause encoding a contrary-to-fact conditional (see (23)).

- (23) Era meglio se mangiava un pochino
be.3.SG.IMPF.IND better if eat.3.SG.IMPF.IND a bit.DIM
“But I want to say, it would have been better, if she had eaten a
bit” (ilfamcv15, C-ORAL-ROM)

In (23), not only has the pluperfect subjunctive in the protasis been substituted but also the required conditional form in the apodosis is missing. Both verbs have been conjugated in the imperfect indicative instead. Here, it is again possible to notice the general tendency towards mood and tense harmony which implies the selection of the same morphological variant for both the main and the embedded clause. According to Leblanc (2009, p. 1014), it is a strong force which “[...]

2 Perception and conceptualization of mood variability in folk linguistics

may be a good candidate for a universal rule of spoken language” and certainly influences mood variation to the detriment of the conservative rule when the latter imposes morphological asymmetries to the sentence.

The analysis of the metalinguistic discourse concerning the prescribed uses of the subjunctive has revealed how folk perception of the phenomenon of variability is limited to only a few contexts. While the if-clause is very salient, complement clauses with a variety of governor verbs are rarely the object of debate, especially if they require the present subjunctive. The majority of non-linguists usually do not perceive the indicative as ungrammatical in these syntactic contexts although standard grammars may still prescribe the subjunctive. This indicates that the non-standard morphological variants have, to a certain extent, already been normalized in informal speech although the subjunctive is not completely discarded either. It remains to be established whether the variation in these and other contexts is completely arbitrary or whether some distinctive parameters, like the mood harmony phenomenon are found to constrain it, systematically predicting the selection of a particular variant. A number of grammarians and linguists who acknowledge the variability make claims that it may be semantically motivated (e.g. Farkas, 1992b; Giorgi & Pianesi, 1997; Quer, 2009). These studies will be addressed in the next section.

If some fundamental aspects influence the variation, these cannot be simply reduced to external or sociological factors regarding register or

2 Perception and conceptualization of mood variability in folk linguistics

education level either. Indeed, the analysis of the metalinguistic discourse has revealed that the morphological variation is a horizontal phenomenon, sweeping across all social groups, education levels and ages. In the instances reported, non-linguists were astonished at realizing that doctors commit errors and even language teachers are found to use subjunctives incorrectly or to discard them in favor of other moods, those very same institutional figures in charge of passing down the Italian language and the correct grammar to the next generation (Yahoo, 2010). Also, in controlled registers and prepared speeches of politicians, TV anchors and radio speakers, morphological distortions and non-standard variants are frequently encountered. Inconsistent use of subjunctive morphology is thus the norm. This fact is interpreted as an indicator that the subjunctive is at least in some syntactic contexts an unstable variable. There is no level of society employing subjunctive morphology categorically, nor is the level of education a predictor of a consistent use of subjunctives in all the syntactic and semantic domains prescribed by grammars. Beyond the very broad gender distinction noticed in public perception, and the assumption which links the subjunctive mainly to formal and controlled registers, social or geographic dimensions cannot completely explain and model the variability. In light of these observations, the learning and representation of the subjunctive must be investigated.

Even when intending to produce a subjunctive form, especially a present subjunctive, errors due to the incorrect selection of the affix occur. These errors generate nonce forms that have never been heard

2 Perception and conceptualization of mood variability in folk linguistics

before, recall *corregghi* “you correct”, *scrivino* “they write” etc. found in the political talk and reported in La Sora Cesira (2018), *Il Giornale* (2018) and Fazio (2018) among others. The fact that speakers fail to extend the proper pattern and generate the appropriate form on the fly points to issues in the productivity of subjunctive morphology. They might be caused by its unstable and infrequent occurrence as a result of the mood variation. Note that these complex forms cannot be retrieved directly from memory as they do not exist. Although this type of error is routinely subject to criticism and irony, it still occurs. Again, looking at the number of instances reported in public discourse, the present tense seems to be more affected by this phenomenon. It is possible to conclude that the morphological forms of the subjunctive system do not enjoy equal productivity rates. Also, the syntactic context has a bearing on the selection of the mood.

The last observation concerns the fact that even if the morphological variation cannot simply be reduced to the speech of a particular social group or setting, social meanings are still consciously attached to it in the Italian community. Folk perception attributes the image of a stereotypical speaker to failures in the selection of the correct subjunctive forms: male, politician, uneducated, impostor. In general, the loss of subjunctive morphology is associated with the idea of a decaying culture, of a degrading linguistic competence and a failure in the educational system. The culprit is usually taken to be the computer- or mobile-mediated communication or any technological innovation. Perceiving phenomena of linguistic variability and change as decline

2 Perception and conceptualization of mood variability in folk linguistics

and loss of a linguistic heritage is very common, particularly in norm-cultivating societies such as the Italian society in which the standard language has a great role as a cultural product and is the object of cultural and national pride. Very recent is a public controversy sparked by a letter to the Ministry of Education signed by more than six-hundred university professors and academics, among whom linguists and grammarians of the Accademia della Crusca, historians, sociologists, journalists and high-caliber personalities in Italian cultural life, like Massimo Cacciari and Ilvo Diamanti (Adinolfi, 2017; De Santis & Fiorentino, 2018). The letter, titled “Saper leggere e scrivere: una proposta contro il declino dell’italiano a scuola” (GruppoDiFirenze, 2007), paints a scenario in which the Italian language is decaying. It laments a sharp decline in the linguistic and writing skills of the students who start their university education. In particular, it reports great deficiencies in the syntax, lexicon, and grammar of university students whose writings purportedly exhibit errors typical of third-graders in elementary schools. Some university faculties, in order to tackle the problem, have introduced extracurricular grammar courses, according to the text. The signers call for urgent ministerial interventions targeting primary education. The appeal coming from the academy triggered a great mobilization of Italian linguists among whom many members of the Italian Linguistic Society (SLI) (De Santis, 2017). In a letter to the Ministry (Lo Duca, 2017), published in newspapers, they review the claims according to which linguistic knowledge is declining and criticize the lack of scientific basis for such

2 Perception and conceptualization of mood variability in folk linguistics

pronouncements. Linguists, such as Luca Serianni and Lorenzo Renzi, authors of grammars of Italian (Serianni, 1989; Renzi et al., 1991), advocate for overcoming the perception of decline and the stereotypes associated with the knowledge of grammar, that are usually exploited to further power dynamics, in favor of the idea that language is dynamic, it changes reflecting our changing reality (Notizie della Scuola, 2017; De Santis, 2017). Although some of its members had signed the original complaint, the linguistic academy, Accademia della Crusca, did not take an official position on the controversy, stressing the importance of further studying the issue. The incident highlights how the problem of reconciling the perception of correctness and the ideal of a “good language” with the currents of variation and innovation does not concern laypeople only, but it involves scholars and linguistic specialists.

The morphological shift which would entail the complete loss of the subjunctive mood in favor of other systems, such as the indicative, the conditional or the imperative, has not taken place to the point that the new forms are completely accepted. There is still pressure to use the correct subjunctive forms at least in more formal contexts. Crucially some people are more aware or feel more deeply about the issue than others. The tension between normativism and praxis is felt within the community at large and generates a negotiation around linguistic practices that the speakers more or less consciously navigate.

3 Semantic accounts of the subjunctive mood in the Romance languages

The interest that the subjunctive has sparked among non-linguists is perhaps exceeded by the attention it has received in the academia. Linguists and grammarians of different traditions have strived to capture the functions and the morphosyntactic behavior of the subjunctive in the Romance languages. Still, there is an effort to develop a unitary notion of the subjunctive, a core property that encompasses all the crosslinguistic and intralinguistic variability (Prandi, 2002; Quer, 2009). The research has focused on whether a core semantics of the subjunctive exists in the mind of the speakers and whether general rules can be formulated to explain its peculiar morphosyntactic distribution.

3.1 The quest for the “core” meaning of the subjunctive

The assumption that the mood distribution is semantically motivated has a long-standing tradition. Semantic theories have been used to explain the occurrence of the subjunctive in independent clauses. In those contexts, the indicative is reported to occur in declarative sentences, a type of utterance whose illocutionary force is to describe the world. For this reason, they are called word-to-world utterances in the sentence taxonomy developed by Searle and Vanderveken (1985). The subjunctive, in direct contrast to the indicative, encodes non-declarative or non-assertive-like utterances (Bolinger, 1968; Hooper, 1975; Hooper & Terrell, 1974; Haspelmath, 2003), i.e. utterances with

special modalized meanings pertaining, among others, to the deontic and bouletic modality. It licenses a directive, volitive and optative interpretation, and is used for utterances with the illocutionary force of a command (24), (25), of a wish (26), and of a request (27).

- 73

3 Semantic accounts of the subjunctive mood in the Romance languages

- (27) Non vi sembri
Neg you.DAT seem.3.SG.PRS.SUBJ

una sottigliezza di blasone!

a subtelty of title

“You must not see it as a subtelty of title” (inatte01, C-ORAL-ROM)

These types of sentences can be loosely categorized as world-to-word utterances: their objective is to get the world to fit their propositional content (Searle & Vanderveken, 1985; Farkas, 1992b, p. 77). Additionally, in independent predication the subjunctive is reported to convey exclamations ((28) and (29)) and encodes the epistemic value of uncertainty in questions (30). In this sense, the mood is conceptualized as a marker of modality. It generates particular interpretations regarding how the truth of the proposition is evaluated against the actual context (Jespersen, 1924) or the background context (Giorgi & Pianesi, 1997) known as the modal base.

- (28) Tu sentissi come cantava
You hear.2.SG.IMPF.SUBJ how sing.3.SG.IMPF.IND
“That boy sang really beautifully!”(ifamnn01, C-ORAL-ROM)

- (29) Avessi visto che giovane! Orca!
Have.2.SG.IMPF.SUBJ see.PP What young! INTERJ

“What a young man! How well-built!” (il fammn01, C-ORAL-ROM)

- (30) Che ci abbia mentito sulla sua identità?
That us have.3.SG.PRS.SUBJ lie.PP on her identity?
“May she have lied concerning her identity?”(il fammn01, C-
ORAL-ROM)

3.1.1 Different surface realizations of the subjunctive

The subjunctive is also and to a greater extent, if one looks at its distribution, the mood of subordination. It occurs in argument, relative and adjunct clauses like conditional or purpose clauses. In argument clauses, it is governed by an element in the independent utterance. This can be a verb, a predicative nominal, a predicative adjective etc. and it is usually referred to as the matrix or the governor. In other types of embedded utterances, the subjunctive is usually licensed by a subordinating conjunction, a pronoun, or an adjective. Grammarians have focused on identifying all the lexical and syntactic contexts where the subjunctive mood is possible. The effort usually results in listing all the governors and lexical elements that license the subjunctive mood in the embedded clause. Nevertheless, categorizations of these elements based on their shared semantics are made. The fact that such generalizations are possible, according to Farkas (1992b, pp. 72-74) reveals that the mood distribution is not random but, indeed, semantically motivated. The assumption behind this argument is that

3 Semantic accounts of the subjunctive mood in the Romance languages

the occurrence of the mood in subordination depends on the meaning of the verb or of the lexical element that licenses it. Additionally, semantic accounts imply that the subjunctive independently conveys a meaning which is in harmony with the lexical semantics encoded in the matrix. For this reason, also in the context of subordination grammarians talk about different types of subjunctives. The concepts are more or less derived from the classification employed for the subjunctive in Latin and do not considerably diverge from the semantic classes used to illustrate the distribution of the mood in independent clauses. However, new categories are introduced, like the factive and emotive class or the subjunctive of polarity. In the following, I review the main categories discussed in the literature:

(1) *directive or jussive subjunctive*

It is the subjunctive governed by verbs of communication like *dire* “say” in utterances with the illocutionary force of a command or by verbs that themselves encode an order or a request (*ask, order, request*). Together with the negated subjunctive it encodes prohibitions.

- (31) Ordino che cessi
Order.1.SG.PRS.IND that cease.3.SG.PRS.SUBJ

la repressione.
the repression

3 *Semantic accounts of the subjunctive mood in the Romance languages*

“I order that the repression be brought to an end”
(lists.peacelink.it, ItWac, Baroni et al., 2009)

(32) I sindacati chiedono che si
The associations ask.INF that REFL.PRO

attivino misure per
activate.3.PL.PRS.SUBJ measures to

dare certezze occupazionali
give.INF certainties working

“The trade unions ask measures to give working assurances”
(telepa.it, ItWac, Baroni et al., 2009)

(2) *optative or volitive subjunctive*

It is licensed by desiderative verbs encoding a want or a wish that the event of the embedded proposition comes to pass (Renzi et al., 1991, p. 416).

(33) Voleva che pulissi
Want.3.SG.PRS.IND that clean.1.SG.IMPF.SUBJ

tutto il giardino [...]
all the garden [...]

3 Semantic accounts of the subjunctive mood in the Romance languages

“She wanted me to clean the entire garden” (ipubdl03, C-ORAL-ROM)

(34) Speriamo che vi portino
hope.PRS.IND that you.PL.DAT bring.3.PL.PRS.SUBJ

davvero molto lontano
really very far

“Let us hope that they bring you very far”(imedts05, C-ORAL-ROM)

(3) *epistemic subjunctive*

It is governed by verbs that express a weaker epistemic commitment with regard to the proposition in the complement clause (Farkas, 1992b, p. 71); in Italian, in particular, contrary to Spanish, French and Romanian, verbs like *believe*, *imagine*, *think*, *admit*, *suppose*, which convey hypotheses, doubts, beliefs, etc. also take the subjunctive if they appear in an affirmative context in contrast to verbs that encode a high degree of epistemic commitment to the truth of the proposition, like *know*, *be certain of* and declarative verbs.

(35) Questo carrozziere a Livorno mi
This mechanic in Livorno me.DAT

3 Semantic accounts of the subjunctive mood in the Romance languages

pare si chiami Carboni
 seem.3.SG.PRS.IND REFL.PRO call.3.SG.PRS.SUBJ Carboni
 “I believe that this mechanic in Livorno is called
 Carboni”(imedts01, C-ORAL-ROM)

(36) [...] non mi sembrava che
 [...] NEG me seem.3.SG.IMPF.IND that

avesse tanto dolore
 have.3.SG.IMPF.SUBJ much pain
 “It didn’t seem to me that he really was in pain”(ifamd110, C-
 ORAL-ROM)

(4) subjunctive of polarity

In the Romance languages, there are contexts, mainly with epistemic verbs like *believe*, *think*, *seem* etc., where the negation of the matrix affects the indicative/subjunctive distribution. In such cases, the negation is taken to affect the epistemic commitment to the truth of the proposition. It conveys non-realistic, improbable eventualities. As such, it prompts the selection of the subjunctive. A positive epistemic commitment, found in affirmed epistemic verb matrices, is related, on the other hand, to the selection of the indicative (Bolinger, 1974; Farkas, 1992b, p. 71; De Mulder & Lamiroy, 2012, p. 210). The negation is also associated with presupposed and non-asserted events, to

3 Semantic accounts of the subjunctive mood in the Romance languages

which the speaker does not commit. They are therefore marked with the subjunctive and contrast to asserted events conveyed with the indicative (Givón 1978; Chafe, 1995; Thompson, 1998).

- (37) No creo que haya
 NEG believe.1.SG.PRS.IND that have.3.SG.PRS.SUBJ

fracasado todavía

break.PP yet

“I don’t think it has broken yet” (19-07-30 ES, Corpus del Español, NOW)

- (38) Creo que esa es la clave
 Belive.1.SG.PRS.IND that this be.3.SG.PRS.IND the key

de los éxitos

of the successes

“I believe that this is the key to success” (19-07-31 ES, Corpus del Español, NOW)

- (39) Non è che la polizia
 NEG be.3.SG.PRS.IND that the police

possa diventare [...]

can.3.SG.PRS.SUBJ become.PP [...]

3 *Semantic accounts of the subjunctive mood in the Romance languages*

“It is not that the police can become [...]” (diario.it, ItWac, Baroni et al., 2009)

(5) *factive/emotive subjunctive*

This category has been created in order to account for the subjunctive in embedded utterances governed by matrices that do not seem to modulate the truth of the event. The selection of the subjunctive in such contexts runs contrary to the core function usually attributed to the mood. Indeed, these utterances do not encode a weak epistemic commitment or irrealis contexts. Rather, the truth of the proposition is assumed by the matrix subject and often an evaluation concerning the event asserted in the embedded proposition is expressed. For this reason, scholars also talk of a *subjunctive of evaluation*. Verbs or predicative adjectives that take the subjunctive are the equivalents of *be happy*, *be surprised*, *be sad*, *be important*, *be better*, *be interesting*, *regret*, *despise*, *rejoice* etc. Renzi et al. (1991) describe this type of subjunctive as follows:

Il congiuntivo può essere usato anche in frasi che non hanno alcun valore dubitativo [...]. Non viene comunicato né che qualcuno ritenga vera la frase dipendente, né che qualcuno dubiti della sua verità. Si tratta piuttosto del fatto che la verità della frase dipendente è presupposta dal parlante. Questa presupposizione è condizione necessaria per frasi di questo tipo: ci si può dolere o rallegrare solo di ciò che si ritiene sia un fatto. [The subjunctive can also be

3 Semantic accounts of the subjunctive mood in the Romance languages

used with utterances that do not convey the meaning of uncertainty. It is neither used to communicate that the speaker believes the dependent clause to be true, nor that they doubt its truthfulness. Rather the truthfulness of the dependent clause is presupposed by the speaker. This presupposition is a necessary condition in these types of utterances. One can only be sad or rejoice about something that one believes to be a fact.] (pp. 418-419)

(40) È interessante che si stia

Be interesting that PRO be.3.SG.PRS.SUBJ

risvegliando il desiderio

awaken the desire

“It is interesting that the desire is rekindling”

(viaggi.repubblica.it, itWac Baroni et al., 2009)

(41) Me alegro que haya

Me rejoice.1.SG.PRS.IND that have1.SG.PRS.SUBJ

personas [...]

people [...]

“I am happy that there are people [...]”(19-0-7-24,
Corpus del Español, NOW)

3 *Semantic accounts of the subjunctive mood in the Romance languages*

It is possible to notice how these classifications mainly have the distribution of the subjunctive in argument clauses governed by verbs in mind. Other types of embedded utterances, like adjunct clauses, have received less systematic categorization. Indeed, they pose some issues for the semantic categories developed as they seem to display some degree of arbitrariness. For instance, adjunct clauses with a temporal value governed by the adverbial locution *prima che* “before” select the subjunctive in Italian. On the other hand, the locution *dopo che* “after” in such utterances selects the indicative. It is difficult to reconcile this contrast with a purely semantically-driven theory.

3.1.2 A binary opposition between the subjunctive and the indicative

In spite of listing a number of semantic domains where the subjunctive operates in subordination, also for embedded utterances, semanticists have been concerned with developing a core and unifying theory of the subjunctive. One of the most influential accounts on subjunctive morphology comes from the generativist work of Giorgi and Pianesi (1997). They argued, building on traditional accounts of mood distribution, that the subjunctive and indicative mood are essentially in binary opposition and a theory that explains the subjunctive should develop two categories in order to account for its relation with the indicative. This idea was at the center of the traditional view which sees the subjunctive and the indicative encoding the dichotomy between *realis* and *irrealis* contexts.

3 Semantic accounts of the subjunctive mood in the Romance languages

The theory proposed by Giorgi and Pianesi (1997), drawing from the work in modal logic (Kratzer et al., 1981) and from Jespersen's (1924) and Farkas' (1992b) conceptualization of notional mood, sees the subjunctive and the indicative as functional categories closely related to modality. The category of mood introduces a set of possible worlds, $W_{(P)}$, or contexts of evaluations, in relation to which an utterance is classified. The worlds are ordered and evaluated based on a number of semantic parameters according to their proximity or similarity to the real world, $W_{(R)}$, which is constituted by the assertions that are known by the discourse participants. These assertions, or declarative utterances, are part of the discourse common ground. The possible worlds, $W_{(P)}$, that are considered similar to the real world, $W_{(R)}$, license the indicative, whereas those that are removed from the actual world license the subjunctive.

Farkas (1992b, p. 74) however, as also Giorgi and Pianesi (1997) themselves, already point out that a comprehensive understanding of the functions and the domains of application of the subjunctive requires more than a binary taxonomy, which has the disadvantage of putting the mood only in relation to the indicative. Such an approach is reductive because the subjunctive interacts with other moods, i.e. the infinitive, the conditional and the imperative, and their related constructions. Limiting the observation to the indicative/subjunctive distribution means limiting the understanding of the syntax and semantics of the subjunctive as well. Farkas (1992b, pp. 72-74) develops a broader taxonomy that summarizes the behavior of the

3 Semantic accounts of the subjunctive mood in the Romance languages

subjunctive in subordination in three major cases: (a) the mood is the result of the semantics of the clause. In such instances, a switch in mood operates a shift in meaning; negation, for one, affects the selection of the mood (polarity subjunctive); (b) the mood is lexically selected by the matrix. Here the semantics of the sentence, like the presence of the negation, does not affect mood choice. Nevertheless, she argues that a semantic factor still operates because the meaning of the mood is always in harmony with the meaning of the matrix (this generates redundancy). For instance, predicates expressing certainty also govern the indicative, predicates expressing uncertainty lexically select the subjunctive; finally, in (c) she acknowledges instances of complete arbitrariness in the mood distribution. However, according to her, they do not defeat a semantic account:

[...] as long as the confines of the arbitrariness one allows are precisely delimited, and as long as the cases that are claimed to be arbitrary do indeed bear the marks of arbitrariness, namely instability of various types: historical, crosslinguistic, dialectal, idiolectal. The claim that a particular phenomenon is semantically driven does not lose its predictive power in the face of exceptions if one can predict where exceptions may or may not occur and how the exceptional cases will behave historically and cross-linguistically (p. 72).

Case (a) in the taxonomy is the only one in which the subjunctive can really be said to have full semantics: the meaning of the sentence

3 Semantic accounts of the subjunctive mood in the Romance languages

crucially changes as a function of the mood selected. Farkas (1992b, p. 70) takes an example from Romanian, reported here in (42). The verb *a spune* “to say” in Romanian appears with the indicative in a declarative context. The subjunctive on the other hand introduces a reported directive.

(42)

- a. Ion a spus [că Maria a plecat].
I. has said that M. has.IND left
- b. Ion a spus [ca Maria să piece imediat].
I. has said that M. SUBJ leave immediately (Farkas, 1992b, p.

70)

Subjunctive forms governed by the equivalents of *say* or other verbs of communication with a directive interpretation are present in other Romance languages. In (43), I report an example from Spanish taken from the Corpus del Español NOW (Davis, 2018).

- (43) Le dije que hablara
PRO.DAT tell.1.SG.PST.IND that talk.3.SG.IMPF.SUBJ

con Pablo y le convenciera
with Pablo and PRO.DAT convince.3.SG.IMPF.SUBJ
“I told her/him to talk to Pablo and to convince him.” (ES19-07-30)

3 Semantic accounts of the subjunctive mood in the Romance languages

Reported imperatives are also possible in Italian. However, Bianchi (2001) notes that such uses are marginal (also in Kempchinsky, 2009, p. 1797). In Italian, as in French, the subjunctive governed by communication verbs like *dire* may also act as an evidential (Giorgi & Pianesi, 2002, p. 205). The distinction must be derived by additional semantic cues, such as the communicative context and the lexical semantics of the embedded verb. In (44) I give an example taken from the C-ORAL-ROM corpus (Cresti & Moneglia, 2005).

- (44) Gli amici dicono che si chiama
 det friends say that REFL.PRO call.3.SG.PRS.SUBJ

 Schen, abbia trentaquattro anni.
 Schen have.3.SG.PRS.SUBJ thirty-four years

 e si prostituisca
 and REFL.PRO prostitute.3.SG.PRS.SUBJ
 “Friends say that her name is Schen, she is thirty-four years old
 and is a prostitute.” (C-ORAL-ROM, Cresti & Moneglia, 2005)

It must be acknowledged, however, that cases such as in (a) where a switch in mood form determines a switch in the interpretation of the entire sentence are on the whole very few. This observation parallels evidence coming from attrition studies, among others in language

3 Semantic accounts of the subjunctive mood in the Romance languages

contact, according to which the subjunctive is lost first in contexts where its occurrence is semantically motivated (Silva-Corvalán, 1994b, p. 269; Ocampo, 1990; Poplack, 1990; Montrul, 2009). If the subjunctive is subject to desemanticization, to the point that its functions become unclear to the speakers, it would make sense that it is discarded first in the contexts where the mood independently contributes to the semantics of the utterance with a modal value that cannot be gathered from the other elements of the sentence.

The reported imperatives with a subjunctive form governed by communication verbs may be one solid instance of a subjunctive with an independent semantic value. However, as hinted previously, they are very rarely used in speech in Italian. No such instances were found in the Italian C-ORAL-ROM corpus (Cresti et al., 2002). On the other hand, subjunctive forms governed by communication verbs in reported declaratives as in (44) still occur in Italian and French. The semantic-based accounts assume that the selection of the subjunctive in reported declaratives anchors the complement proposition to another source, the world of the subject, $W_{R(Su)}$, rather than the world of the speaker itself, $W_{R(Sp)}$ (Speas, 2004). As such, it does not commit the speaker to either a positive or a negative evaluation of the proposition and the proposition is not taken to enter the common ground of the discourse context, i.e. the set of propositions that are known to be true in the actual world: it indexes “afactuality” (Farkas, 1992b, p. 81). It is however a matter of debate whether nowadays the subjunctive in lieu of an indicative actually generates this evidential interpretation as it remains

an open question whether the mood has kept its original semantic value in that particular syntactic context or not. The study conducted by Montrul (2007) on Spanish heritage speakers shows how speakers fail to identify a shift in the semantic interpretation of the sentence produced by the subjunctive/indicative contrast.

3.1.3 Prototypical subjunctive verbs

I leave aside the analysis of the mood distribution in the particular contexts where the matrix can take both forms. Far more frequent, as corpus data suggests (Poplack et al., 2013, 2018; Porto Dapena, 1991), are occurrences of subjunctive morphology of the type (b). The mood is largely licensed by the matrix of argument propositions (lexical conditioning). Among many different types of matrices that trigger the occurrence of the mood, predicates and lexical elements encoding the directive, volitive and optative modality like the equivalents of *want*, *wish*, *order*, *request* etc. take the subjunctive in all Romance languages. They are claimed to be the locus where the subjunctive mood is most stable crosslinguistically (see Kempchinsky, 2009). Additionally, according to the semantic view, they convey the core function of the subjunctive: encoding non-declarative, *irrealis* contexts that do not commit the subject to the truth of the propositions. In Farkas' (1992b, p. 89) terms, based on the work on modal logic by Kratzer (1977, 1979) and Kratzer et al. (1981), the predicates introduce a set of possible worlds, $W_{(p)}$, according to which the argument proposition is evaluated (modal anchor). These worlds are regarded to be distant enough from

3 Semantic accounts of the subjunctive mood in the Romance languages

the real world, $W_{(R)}$, constituted by the common ground of assertions, and therefore they require the subjunctive.

However, the analysis of natural data reveals that even those predicates never condition the occurrence of the subjunctive categorically. They are found to select the indicative as well. The actual rates of occurrence of indicatives differ across the Romance languages (see chapter 5). An example is reported in sentences (45a) and (45b) taken from the C-ORAL-ROM corpus.

(45)

a. Voleva pulissi tutto il giardino davanti.

“He wanted me to clean (IMPF SUBJ) the entire garden.”

(ipubdl03, C-ORAL-ROM)

b. Io volevo che ci raccontavano questa cosa. (ifamcv17, C-ORAL-ROM)

“I wanted them to tell (IMPF IND) us this thing.”

In these contexts, the morphological alternation between the indicative and the subjunctive does not generate any apparent shift in the semantic interpretation of the sentence. The illocutionary force of the embedded utterance is in fact provided by the matrix element and not by the embedded mood. In (45a) and (45b) it is the matrix verb *volere* “want” that generates the modal reading: the event in the argument clause is a non-actual state wished for by the referent of the subject. Thus, it is not the case that the arbitrariness is an entirely separate phenomenon from lexical conditioning as Farkas’ (1992b) classification suggests.

3 Semantic accounts of the subjunctive mood in the Romance languages

Although Farkas (1992b) already acknowledges lexical conditioning and arbitrariness in the selection of the subjunctive, ultimately she, like Giorgi and Pianesi (1997), does not give up on the idea that the mood has a core semantics also in the embedded predication. According to her, the uses that embody this core semantics should be the most stable crosslinguistically and intralinguistically. This position is shared by different scholars in the semantic tradition (see Kempchinsky [2009] for a similar account): across languages there will be least morphological variation in the contexts that encode the core functions attributed to the subjunctive.

The assumption that a core meaning or a core set of functions exists for the subjunctive which can explain its complex morphosyntactic behavior has stimulated an effort to account for the mood distribution in contexts that would seem to defy the general understanding of the mood category as an index of modality.

3.1.4 Two distinct types of subjunctive

While maintaining the position that the occurrence of the subjunctive can be explained by a semantic theory, not all scholars agree that the main functions of the subjunctive can be subsumed under a unitary theory. According to Stowell (1993) and Quer (1997), it is not possible to treat the subjunctive as a uniform morphological marker, and scholars should give up on an all-encompassing definition. They argue for the existence in Romance of two formally distinct types of subjunctive: “intentional subjunctive (e.g. triggered by an intentional

verb like 'want') and polarity subjunctive (e.g. licensed by a matrix negation or a question operator)" (Quer, 1997, p. 171). It is important to note that this classification is not exclusively based on the semantic-functional properties of the larger context where the mood occurs. It also hinges on its morphosyntactic behavior. Four main features that distinguish the subjunctive of negation from the intentional subjunctive are identified by Quer (1997). They are reported in the following:

- (i) The intentional subjunctive cannot encode all temporal specifications and it is rather considered to anaphorically refer to the tense of the indicative in the matrix clause (Picallo, 1985). Generally, a past subjunctive governed by a present indicative is not possible (see (46b)). This is however possible with the subjunctive governed by negation (47a). The examples are taken from Quer (1997, pp. 172-173) and are based on Catalan:

(46)

a. Volia que
want.3.SG.IMPF.IND that

acabés la tesi
finish.1.SG.IMPF.SUBJ the dissertation
“S/he wanted me to finish the dissertation.”

b. *Vull que
want.1.SG.PRS.IND that

3 *Semantic accounts of the subjunctive mood in the Romance languages*

acabés la tesi
 finish.3.SG.IMPF.SUBJ the dissertation
 “*I want that s/he finished the dissertation”

- (ii) The intentional subjunctive does not alternate with the indicative while the polarity subjunctive does (see (47a) and (47b)).

(47)

a. No recorda que en Miquel
 NEG remember.3.SG.PRS.IND that the Miquel

treballés de nit
 work.3.SG.IMPF.SUBJ of night

b. No recorda que en Miquel
 neg remember. 3.sg.prs.ind that the Miquel

treballa/treballava/treballarà de nit
 work.3.SG.PRS/IMPF/FUT.IND of night

“S/he does not remember that Miquel works/worked/will work at night.”

- (iii) Intentional matrices like *want* govern only their direct complement. The negation matrix however has a larger

scope and can govern the complement in consecutively embedded constructions. The subjunctive is thus possible only in the immediate complement of volitive verbs but appears in consecutive clauses governed by a negation. This is shown in the examples (48) and (49) from Catalan taken from Quer (1997, p. 174).

(48)

- a. Vull [que creguin
want.1.SG.PRS.IND that think.3.SG.PRS.SUBJ

[que ens agrada]]
that us please.3.SG.PRS.IND
“I want them to think we like it.”

- b. *Vull [que creguin [que
I-want that think.3.SG.PRS.SUBJ that

ens agradi/agradés/hagi agradat]]
us please.3.SG.PRS/IMPF/PP.SUBJ

(49)

- a. No crec [que pensi
not believe.1.SG that think.3.SG.PRS.SUBJ

[que li convé]]

3 Semantic accounts of the subjunctive mood in the Romance languages

that him favor.3.SG.PRS.IND

“I don't believe s/he thinks it's convenient for him/her.”

b. No crec

not believe.1.SG.PRS.IND

[que	pensi	[que
that	think.3.SG.PRS.SUBJ	that

li	convinguì]]
him	be-convenient.3.SG.PRS.SUBJ

“I don't believe s/he thinks it's convenient for him/her.”

The disjoint reference effect, also known as the obviation effect is a property of complement clauses governed by matrices that select the subjunctive (Kempchinsky, 1987): the subject of the embedded complement clause and the subject of the matrix clause cannot encode the same referent. This effect is found in the volitional subjunctive (see Spanish in (50) taken from Kempchinsky [2009, p. 1789]. On the other hand, coreference is possible with the negation matrix (see (51) taken from Quer [1997, p. 174]).

(50)

a. Yo_i quiero que *pro**_{i/j}
I want that pro

3 Semantic accounts of the subjunctive mood in the Romance languages

baje a la calle.
 go-down.*1SG/3SG.PRS.SUBJ to the street
 “I want that *I go down / that s/he goes down to the street.”

b. [El jefe del departamento]i insistía
 The chair of-DET department insisted

en que *pro*_{*i/j} enseñara
 on that PRO teach.1/*3.SG.IMPF.SUBJ

esa clase.
 that class

“The chair of the department insisted that *he/ I teach that class”

c *Tu siempre desees que *pro*
 *You always desire that pro

salgas primero.
 go-out.2.SG.PRS.SUBJ first

(51)

No crec *pro*_i [que la
 NEG think PRO that her

3 *Semantic accounts of the subjunctive mood in the Romance languages*

convidi *pro*_i]
invite.1.SG.PRS.SUBJ
“I don't think I will invite her.”

Based on the characteristics of the volitive and the polarity subjunctive, they could be incorporated respectively in category (b) for the lexically-selected subjunctive and in category (a) for the subjunctive licensed by the semantics of the sentence proposed by Farkas (1992b) and discussed above.

Concerning feature (ii) describing the mood distribution in the intentional and polarity environments, it was already shown that, in the case of lexical selection, the volitive matrix does not guarantee a categorical occurrence of the subjunctive mood. Instances of variation between the subjunctive and the indicative are found in speech in the Romance languages. Only, this morphological alternation cannot be traced back to two different functions or semantic domains as is assumed to be the case for the polarity subjunctive.

As already hinted above, the polarity subjunctive or the subjunctive licensed by negated epistemic verbs has been taken to embody one of the core functions in the Romance languages, namely the retraction of an assertion. In this functional domain, it has been considered to occur in paradigmatic opposition with the indicative. The latter is the mood of assertion encoding affirmative epistemic commitment. This semantic divide seems to be supported by the indicative/subjunctive distribution

3 Semantic accounts of the subjunctive mood in the Romance languages

under “seem” verbs in Spanish. Quer (2009, p. 1781) reports that both mood forms can occur under *parer* but they generate different interpretations. The alternation is meaningful. The example he provides is given in (52).

(52)

- | | |
|--|------------------------------|
| a. Parece que llueve (Spanish) | y de hecho está lloviendo |
| “It seems that it is raining.IND.” | “and in fact it is raining” |
| b. Parece que llueva | #y de hecho está lloviendo |
| “It looks as if it were raining.SUBJ.” | “#and in fact it is raining” |

In (52a), the indicative marks a positive epistemic commitment towards the event embedded in the complement clause, the fact that it is raining. In (52b), the subjunctive, on the other hand, generates a counterfactual interpretation of the embedded eventuality. The possibility that it is raining is deemed unreal. For this reason, a continuation like *y de hecho está lloviendo* (“and in fact it is raining”) is possible in (52a) but considered pragmatically odd in (52b). (Quer, 2009, p. 1781).

This distinction observed in languages like Catalan, Spanish and Romanian prompts Quer (2009), based on Giannakidou (1997) and Farkas (1992b), to identify a main divide in the mood distribution between epistemic verbs, which, according to him, systematically take the indicative across languages, and volitive or jussive predicates which take the subjunctive. It is unclear how this opposition between volitive and epistemic environments is supported on semantic grounds.

3 Semantic accounts of the subjunctive mood in the Romance languages

However, as is known, Italian represents a major exception in the theorization of a subjunctive of polarity. As a matter of fact, in standard Italian the subjunctive is prescribed and used in complements of epistemic verbs in affirmative contexts (Dardano & Trifone, 1997). Polarity does not affect the mood distribution. Also, in this case, grammarians give a functional explanation which associates the mood to the semantics of the matrix predicates that govern it. The fact that the subjunctive is selected by “belief” verbs like *credere* “believe”, *pensare* “think”, *sembrare* “seem” is explained by presenting the subjunctive as the mood that conveys a subjective representation, a state of mind, in contrast to the objective truth encoded with the indicative. The subjunctive is, in this sense, the mood that encodes personal opinions, doubts and possibilities (Dardano & Trifone, 1997). We see here a challenge to an all-encompassing theory of mood distribution coming from crosslinguistic variability. It is also possible to notice how, by shifting the focus to certain features of the predicates, a semantic theory can *a posteriori* formulate a semantic divide that best fits the data at hand. As a matter of fact, Giorgi and Pianesi (1997), in order to solve the problem posed by this diverging pattern and generalize over languages, propose that different languages arbitrarily grammaticalize the distinction between subjunctive and indicative along a continuum. Belief verbs in Italian are taken to encode possible worlds that are far removed from the actual one (non-realistic environments). For this reason, they select the subjunctive. In the other Romance languages, such predicates encode contexts of evaluation that are considered closer

3 Semantic accounts of the subjunctive mood in the Romance languages

to the real world (weakly realistic environments) and they thus govern the indicative. Also, in this case, it is unclear how this semantic characterization can be specified within the predicate itself. This assumption entails that Italian speakers actually conceptualize belief verbs differently from speakers of other Romance languages.

Unsurprisingly, morphological variation is found in argument clauses governed by belief verbs in Italian. The phenomenon is so extensive in speech that grammars also acknowledge it as a new norm (Telve, 2011). In (53a) and (53b), the matrix predicate *sembrare* “seem” governs the indicative and the subjunctive respectively. The same type of mood variability is displayed in (54a) and (54b) under matrix predicate *credere* “believe”.

(53)

- a. Mi sembra che ha ritrovato un po’ un equilibrio.
“It seems that she has found (IND) a bit of balance.” (ilfamcv01, C-ORAL-ROM)
- b. Sembra che sia imbarazzato ogni volta che dice qualcosa.
“It seems that she is (SUBJ) embarrassed every time that she says something.” (ilfamcv12, C-ORAL-ROM)

(54)

- a. Io credo che tutto ritorna.
“I believe that everything comes back around (IND).”
(imedin04, C-ORAL-ROM)
- b. Il primo incontro credo risalga ai suoi sedici anni.

3 Semantic accounts of the subjunctive mood in the Romance languages

“I believe that the first encounter dates back to (SUBJ) when he was sixteen years old.” (imedin01, C-ORAL-ROM)

Recall that this type of mood alternation in argument clauses governed by positive epistemics is actually claimed to be meaningful in Spanish because it generates an interpretative contrast (Quer, 2009). However, in Italian, no such contrast can be detected. The subjunctive does not license a counterfactual interpretation. To make sure that this is the case, argument clauses governed by an epistemic verb were presented in pairs to thirty native speakers of Italian. The sentence pairs differed only in the mood form of the embedded predicate: an indicative and a subjunctive respectively. An example is reported in (55a) and (55b). All the subjects found that the sentence pairs yielded the same interpretation. Crucially, they could not trace the mood alternation to a gradation in the epistemic commitment of the matrix subject to the event described in the embedded utterance. This is already confirmed by Giorgi and Pianesi (1997, p. 223) who, based on their own intuition, find that the meaning of the sentence is not altered by mood changes.

(55)

- a. Credo che passa lei per l’ufficio domani.
“I believe she will swing (IND) by the office tomorrow.”
- b. Credo che passi lei per l’ufficio domani.
“I believe she will swing (SUBJ) by the office tomorrow.”

3 Semantic accounts of the subjunctive mood in the Romance languages

Farkas (1992b), unlike Giorgi and Pianesi (1997), would account for this contrast between Italian and other Romance languages by positing that the mood is a feature specified in the predicate in Italian. This means that the epistemic predicates lexically select the subjunctive while other languages allow the mood to vary as a function of the interpretative features of the sentence. We see already the tension here between a purely functional account and a theory that involves lexical and syntactic constraints to model the mood distribution. Also, positing a lexical constraint does not shed any light on why the seemingly meaningless indicative/subjunctive alternation can still be encountered in argument clauses governed by these predicates.

It would also be wrong to deduce from the morphosyntax of epistemic verbs in Italian that the subjunctive of polarity does not exist at all in this language. In fact, in Italian there are predicates for which the mood distribution correlates with the presence of the negation. One instance is the verb *ricordare* “recollect /remember” as in (56a) and (56b), or *sapere* “know”.

(56)

- a. Non ricordo che mi abbiano parlato molto di questo.
“I don’t remember that they told (SUBJ) me much about it.”
(eddyburg.it, ItWac Corpus, Baroni et al., 2009)

- b. Ricordo che la mia maestra era appassionatissima di pittura.
“I remember that my teacher was (IND) fond of painting.”
(giuda.it ItWac Corpus, Baroni et al., 2009)

3 Semantic accounts of the subjunctive mood in the Romance languages

This polarity effect can still be established even if the negated matrix does not exclusively select the subjunctive, and the morphological variation is again, as in many other instances, a common phenomenon (see (57a) and (57b)).

(57)

- a. Non ricordo che ci sia stata una battaglia. (montag.it, ItWac Corpus, Baroni et al., 2009)

“I do not recall that there has been (SUBJ) a battle.”

- b. Non ricordo che nome mi ero inventata. (disegno.it, ItWac Corpus, Baroni et al., 2009)

“I do not recall which name I invented (IND).”

Among the features of the subjunctive of polarity analyzed by Quer (1997) in Catalan, he found that, unlike other argument clauses, coreference is indeed possible between the matrix subject and the subject of the embedded subjunctive when the mood is licensed by the negation (feature iv discussed above). This is also true for verbs which encode the semantics of negation like *negar* “deny” and *dubtar* “doubt”. In Italian, however, although coreference is possible in theory, the disjoint reference effect seems to play a role both in the speakers’ interpretation and in their production of such sentences. In sentence (58a) the verb embedded in the argument clause is ambiguous for verb person: the subjunctive present form with inflectional affix *-a* encodes

3 Semantic accounts of the subjunctive mood in the Romance languages

the first, second and third person singular. The subject of the matrix predicate on the other hand, unambiguously encodes the first person singular.

The sentence was presented to native speakers of Italian. When asked to identify the referent of the embedded predicate, some of the participants did not recognize the ambiguity encoded in the embedded predicate and directly assigned a third person referent (*lui/lei* “he/she”) to it, thus eliminating the coreference between the matrix subject and the subject of the embedded clause. Others, while acknowledging the ambiguity (the predicate could refer both to a first-person subject, coreferent with the matrix subject, and to a third-person referent), stated that they would instinctively assign a third-person referent to the embedded verb, suggesting that a first-person referent would be more aptly encoded with a control infinitive clause introduced by the preposition *di*. An example is given in (58b). These judgments reflect the fact that coreference is rarely expressed in subjunctive-selecting clauses⁴. Another construction, the embedded infinitive, is far more commonly used in that context (in case the subject of the main clause and the one of the embedded utterance encode the very same referent).

⁴Additional features influence the possibility of (or correlate with) coreference between the matrix subject and the subject of the embedded subjunctive clause. These have been expounded by Ruwet (1984) and include the presence of modal verbs in the embedded clause and particular temporal/aspectual features of the two clauses. For instance, the absence of *consecutio temporum* between the matrix and the embedded verb such as in (58a) correlates with the possibility of having coreference. These morphosyntactic features will be discussed in more detail in section 3.2.

(58)

a. Non ricordo che *PRO* abbia detto di non poter venire alla riunione

“I don’t remember that I/he/she said (SUBJ) that I/he/she could not come to the meeting”

b. Non ricordo di aver detto di non poter venire alla riunione

“I don’t remember having (INF) said that I cannot come to the meeting”

Also, with the verb *negare* “deny”, which licenses the subjunctive of polarity, coreference should be possible according to Quer (1997). The analysis of corpus data from the ItWac Corpus (Baroni et al., 2009), however, reveals a particular phenomenon. Whenever the matrix subject and the subject of the embedded utterance happen to be co-referents, the indicative is used instead of the subjunctive under *negare*, especially in the first person. No other characteristics of the sentence, whether semantic or syntactic, can account for the systematic selection of the indicative in such cases. An example is already displayed in (57b). Further instances are reported here in (59).

(59)

a. Non nego che ho pianto quando ho dato il nipotino all’assistenza sociale.

“I don’t deny that I cried (IND) when I gave my nephew to social services.” (filmup.leonardo.it, ItWac Corpus, Baroni et al., 2009)

3 Semantic accounts of the subjunctive mood in the Romance languages

- b. Non nego che a volte ho momenti di sconforto.
“I don’t deny that sometimes I have (IND) moments of dejection.” (edscuola.it, ItWac Corpus, Baroni et al., 2009)
- c. Però non nego che ci soffro davvero tanto
“I don’t deny that I suffer (IND) a lot about it.” (verobusa.etz.itc, ItWac Corpus, Baroni et al., 2009)

We can conclude that there are two constructions which take up the function of conveying coreference between the subjects of the main and the embedded clause:

- (a) The embedded infinitive clause introduced by conjunction *di*.
The latter requires an empty subject controlled by the subject of the matrix (see (60)) (Giorgi and Pianesi, 2004a, p. 194).

(60) Marco crede di *pro* aver preso peso.

Marco believes to *pro* have.inf gained weight.

“Marco believes that he has gained weight.”

- (b) The argument clause introduced by the complementizer *che* and the embedded verb marked for the indicative rather than the subjunctive (examples (59a) and (59b), and (59c) above).

3.1.5 Weaknesses of the semantic account

The evidence coming from natural data and reported so far suggests that the mood distribution cannot be properly understood by resorting exclusively to a semantic theory. Nor is the morphological variation completely random either. Additional factors influence the mood distribution, like the need to override the subjunctive disjoint reference effect in order to encode the same referent in both the main and the embedded utterance, which triggers the selection of the indicative in lieu of the prescribed subjunctive.

Also, natural language data suggests that indicative clauses are not the only syntactic constructions that concur with the subjunctive to encode certain functions. We see in (58b) how infinitive clauses fulfill functions for which the subjunctive, due to its morphosyntactic characteristics, does not represent the best option available to the speaker. Thus, the space of morphological variation that must be mapped does not involve the indicative/subjunctive alternation alone. It also concerns the occurrence of other moods, like the infinitive and, as it will be discussed in the following chapter, the conditional and the imperative, with which the subjunctive shares functional domains of application.

Finally, the evidence reported indicates that a strict divide between two core functions of the subjunctive, polarity and volitive, as proposed in Quer (1997) and further expanded in Quer (2009), cannot be straightforwardly supported either on morphosyntactic or on semantic grounds from a crosslinguistic perspective.

3 Semantic accounts of the subjunctive mood in the Romance languages

According to Quer's theory, the disjoint reference sets apart the volitive subjunctive from the subjunctive of polarity, the latter should permit coreference. However, corpus data shows how the disjoint reference effect impacts the interpretation of the sentence and the selection of the mood also in contexts where the subjunctive is licensed by the negation operator.

Additionally, polarity or the retraction of an assertion, does not impact mood selection equally in the Romance languages. In Italian, the negation does not affect mood distribution in the contexts outlined by a theory based on the epistemic commitment of the subject, namely with the verbs of the epistemic class. However, negation does, to a certain extent license the subjunctive with other verb matrices, like *ricordare* "remember" (see (57a)), *sapere* "know" (see (61)), *capire* "understand" (62), or or *essere* "be" in the locution *non è che* (see (63)).

(61) Non sapevo che tu avessi la passione per la fotografia.

"I didn't know that you have (SUBJ) a passion for photography"
(ilfamcv03, C-ORAL-ROM)

(62) Però non riesco a capire cosa significhi questo.

"However, I don't get what this means (SUBJ)." (ilfamdl113, C-ORAL-ROM)

(63) È ovvio che questo non è che aiuti molto i semigiovani ad avvicinarsi al mondo master.

3 Semantic accounts of the subjunctive mood in the Romance languages

“It is obvious that this does not help (SUBJ) the semi-young get closer to the world of master education.” (atleticanet.it, ItWac Corpus, Baroni et al., 2009)

Also, in the other Romance languages where negation is assumed to play a major role in mood distribution, natural data coming from corpus studies indicates that negation does not systematically license the subjunctive mood. The subjunctive governed by negated epistemic verbs is virtually absent in Portuguese, it occurs very infrequently in French (only 13% of the times) and not always in Spanish (77%) (Poplack et al., 2018, p. 17).

The semantic investigations on the notion of mood in the Romance languages have identified another challenge to an all-encompassing theory of the subjunctive. This challenge comes from factive-emotive and evaluative predicates. As Giorgi and Pianesi (1997, p. 218) aptly summarize, factive-emotive verbs encode an emotional state generated by the event described in the complement clause (*be happy, regret, rejoice, be relieved* etc.); evaluative verbs encode an evaluation of the embedded event (*be weird, be evident, be natural* etc.)

This class generally licenses the subjunctive in Romance languages (see (40) and (41)). An exception however is constituted by Romanian where the indicative is governed by these verbs, as is also the case for other languages in the Balkans (Farkas, 1992b, p. 71). The problem is constituted by the fact that these predicates, unlike other subjunctive-selecting matrices, do not modalize the truth of the embedded

3 Semantic accounts of the subjunctive mood in the Romance languages

proposition. The model of evaluation for the embedded event remains the actual world, $W_{(R)}$, and not a set of possible worlds $W_{(P)}$.

In (64), the fact that the addressee could not turn in her work is an event taken to have occurred in the real world.

(64) *Mi dispiace che tu non abbia fatto in tempo a consegnare il tuo lavoro.*

“I am sorry that you could (SUBJ) not turn in your work in time.”

Farkas (1992b) notes that this type of predicate did not govern the subjunctive in the past. The mood is a result of a diachronic development of the subjunctive. She finds a solution to this apparently counterintuitive behavior of the subjunctive and to mood variation across languages by suggesting that “[...] the factivity of these predicates is connected to their governing the indicative and their emotive [and evaluative] character is responsible for their governing the subjunctive” (p. 71). More specifically, the factivity lies in the complement clause. The fact that the embedded event is true in the actual world, $W_{(R)}$, makes it possible for the embedded utterance to select the indicative. However, according to Farkas (1992b, p. 101), the matrix verb, which denotes either an evaluation or an emotional state, does not introduce any particular world (afactuality), a trait that places this type of verb nearer to desideratives, directives and modals and makes the subjunctive a viable option.

3 *Semantic accounts of the subjunctive mood in the Romance languages*

A different explanation given by semanticists and briefly introduced above draws on the declarative/non-declarative opposition encoded in the indicative/subjunctive divide. Here, the subjunctive is taken to encode the fact that the embedded proposition is not asserted but rather presupposed by the speaker. The presupposed event is, according to Quer (2009, p. 1782), integrated into the common ground of the conversational context. He finds support for this theory in verbs like “admit” which in Italian and French license the subjunctive. The embedded event is considered true. However, the subjunctive signals that the embedded utterance is taken to become part of the common ground (see (65) and (66a)). The latter example is extracted from Grevisse (1993, p. 1618) and reported in Quer (2009, p. 1782). An indicative in such context would signify according to Quer that the embedded proposition is simply asserted by the speaker (see (66b) from Quer [2009, p. 1782]). His definition seems to draw on the traditional binary functional distinction between foregrounding and backgrounding in discourse analysis. Old or presupposed information is marked with the subjunctive and it is thus backgrounded. The indicative, on the other hand, signals new asserted information. The embedded event becomes foregrounded (Givón, 1987). In (66b) the complement, i.e. the fact that the addressee of the locution is right, is foregrounded with the indicative. On the other hand in (66a) the action encoded in the matrix verb *admettre*, the act of conceding, is part of the conversational background.

(65) Ammetto che il mio post fosse impulsivo e poco istruttivo.

“I admit that my post was (SUBJ) impulsive and not informative.”
(centomovimenti.it, ItWac)

(66)

a. J’admets que vous ayez raison

“I admit that you are (SUBJ) right”

b. J’admets que vous avez raison

“I admit that you are (IND) right”

The presupposition/assertion dichotomy has been used by Quer (2009, p. 1782) among others to explain the selection of the subjunctive also in adjunct embedded clauses whose meaning does not depend on the semantics of a matrix predicate. They are the concessive clauses introduced by conjunctions like “although”, but also temporal clauses. Indeed, Quer points out that in Catalan and Spanish the temporal clause equivalent of “after” selects the subjunctive. However, in Italian, as already hinted above, the temporal clause introduced by the equivalent of “after”, *dopo che*, usually selects the indicative. The subjunctive is considered possible by grammarians only when the locution refers to facts that have not taken place yet and are seen as a future eventuality (see (67)). The temporal clause introduced by *prima che* “before”, on the other hand, selects the subjunctive, according to the standard norm, also when it refers to an event that has occurred in the past in relation to the time of the utterance (see (68)).

3 Semantic accounts of the subjunctive mood in the Romance languages

Even when considering only the language norm, it is unclear how this idiosyncratic behavior of the mood across languages can be explained by only drawing on presupposition.

(67) La prima revisione può aver luogo solo dopo che sia trascorso un anno.

“The first exam can take place only after a year has passed (SUBJ).”
(ispesl.it, ItWac)

(68) È riuscita ben prima che nascessero i Verdi.

“She succeeded, way before the Green Party was founded”
(eddyburg.it, ItWac)

At this point, we might also point to notional inconsistencies across semantic theories in the literature, which generate confusion for the reader. It was shown how, according to Farkas (1992b), the common ground is constituted by the set of propositions known to be true in the actual world by the discourse participants. The indicative encodes assertions that become part of this common ground, i.e. they are integrated in this set of utterances that classify the actual world, whereas embedded propositions encoded with the subjunctive do not become part of the common ground. They convey either alternative realities (world of dreams, fiction worlds) or possible eventualities $W_{(p)}$. For Quer (2009), on the other hand, the subjunctive has the effect of integrating the embedded event in the common ground, i.e. the

propositions that are true in the actual world. His notion of common ground, however, has to do with presupposed truths rather than asserted ones.

3.1.6 Conclusion

From the examination of the literature on the semantic basis of mood distribution, four main dimensions have come to the foreground in the explanation of the grammar of the subjunctive:

- (a) Modality captured in terms of closeness of a set of possible worlds, $W_{(P)}$, encoded in a proposition to the real world, $W_{(R)}$ according to the speaker, $W_{R(Sp)}$.
- (b) Evidentiality: the act of anchoring the utterance to a source of information other than the speaker. The content of the complement clause is true of $W_{(R)}$ of the subject, $W_{R(Su)}$, but not of the world of the speaker $W_{R(Sp)}$. The mood signals the contrast between these two sources of evaluation.
- (c) Assertion vs. negation
- (d) Assertion vs. presupposition or the foregrounding/backgrounding functional contrast.

These dimensions interact in an unsystematic manner: they often concur to explain the same phenomena. An example is the subjunctive governed by factive verbs. It has been accounted for by resorting either to the presupposition of truth encoded in these verbs or by highlighting their emotive character. Another relevant case is the polarity

3 Semantic accounts of the subjunctive mood in the Romance languages

subjunctive licensed by negated epistemic verbs. It has been explained either by claiming that the subjunctive encodes the retraction of an assertion or by invoking modality: negated epistemics unlike asserted ones encode possible worlds that are more removed from the actual world and thus require the subjunctive. Other times, one of the dimensions becomes relevant to explain the mood distribution in some languages rather than others. Recall in this regard the *reportive subjunctive* under “say” verbs operating in some Romance languages but not in others, or the *presupposition subjunctive* which might fit the Spanish and Catalan data better than the Italian data.

As has been shown, the theory struggles to capture the normative use of the subjunctive. In addition to this, it is evident that the data coming from natural speech does not neatly pattern with any of the semantic theories reviewed. No matter how refined, they are unable to account for the non-normative and seemingly idiosyncratic mood alternation which takes place also in controlled speech, as amply reviewed in chapter 3.

Semantic theories conceptualize an abstract idea of mood which lends itself well to and is amply used for instructional purposes. However, with the increasing modification that the mood systems are undergoing in speech (Silva-Corvalán, 1994b), theories of language change but also of language learning could best account for the complex phenomena concerning the mood distribution in speech, by examining how the system develops diachronically, how it is passed down to the next generations and how the new generation of learners constructs

their morphological knowledge based on the data available to them through the experience of natural language.

3.2 The subjunctive mood at the interface between syntax and semantics

The semantic properties of the subjunctive have proven to be elusive and difficult to integrate into a coherent unitary theory. A promising approach has been the examination of the subjunctive's peculiar morphosyntactic characteristics and their interaction with the semantics of the larger sentence.

The research has focused for the most part on three main properties:

- (a) The temporal interpretation of subjunctive clauses
- (b) Obviation strategies and the subjunctive disjoint reference constraint
- (c) The behavior of the complementizer of subjunctive-selecting argument clauses

Some of them have already been briefly introduced above. In the following, based on the review of previous work, I will closely examine each property and highlight how the three properties interact with each other.

3.2.1 Temporal characteristics of the subjunctive

It has been argued that the subjunctive in embedded clauses is inherently tenseless and that it exhibits anaphoric tense: it cannot encode temporal information independently of the matrix clause and it inherits the temporal features of the main clause (Picallo, 1984). As such, it is governed by a strict *consecutio temporum*, i.e. a temporal agreement. The present in the main clause selects the present in the embedded locution whereas the past triggers a past form (Giorgi, 2009, p. 1842). This is shown in the examples provided by Giorgi (2009, p. 1842) and reported here in (69).

(69)

- a. Gianni crede che Maria sia/ *fosse incinta.
'Gianni believes that Maria is (PRS SUBJ)/ *was (PST SUBJ) pregnant.'
- b. Gianni credeva che Maria fosse/ *sia incinta.
'Gianni believed that Maria was (PST SUBJ) / *is (PRS SUBJ) pregnant.'

According to Giorgi (2009, pp. 1842-1843), additional evidence for the fact that the subjunctive lacks a specific temporal value comes from the temporal adverbs it can select. Unlike the indicative, the past subjunctive governed by certain matrix verbs, for instance the epistemic verbs, can be found with modifiers encoding past, present or future time reference with respect to the speaking time or to the time of the main event. See sentence (70) taken from Giorgi (2009, p. 1842).

(70) Gianni credeva che Maria partisse ieri/ adesso/ domani/ il giorno dopo/ [il giorno prima].

‘Gianni believed that Maria left (PST SUBJ) yesterday/ now/ tomorrow/ the next day/ [the previous day].’

In order to further characterize the difference between the subjunctive and the indicative and outline the rules that govern the sequence of tense (SOT) for these two moods in embedded utterances, scholars have looked at a property called *double access reading* (DAR) (Giorgi & Pianesi, 1997, 2000, 2004b; Higginbotham, 2002; Abusch, 1997). This property holds when the event expressed by the embedded predicate “[...] is evaluated twice: once with respect to the utterance time and once with respect to the time of the main event” (Giorgi, 2009, p. 1838). This is the case for embedded clauses selecting the indicative mood in Italian. An example usually given in order to highlight the DAR property involves the indicative future tense in embedded clauses, such as in (71).

(71) Salvatore ha annunciato che Antonietta tornerà a casa.

‘Salvatore announced that Antonietta will come back (FUT) home.’

The embedded utterance in (71) expresses futurity both in relation to the event in the main clause, the time in which Salvatore’s

3 Semantic accounts of the subjunctive mood in the Romance languages

announcement took place, and in relation to the speaking time. The interpretation that Antonietta might already have come home at the time of speaking is excluded. The subjunctive, on the other hand, does not always instantiate DAR.

(72) Raffaella pensava che i regali arrivassero la settimana successiva.

“Raffaella thought that the presents would arrive (IMPF SUBJ) a week later.”

In (72) the embedded predicate marked for the imperfect subjunctive does not give any temporal indication in relation to the speaker. The embedded event may or may not have taken place at the time of the utterance. Also, the time adverbial *la settimana successiva* (“a week later”) has to be understood in relation to the time of the main event, the time in which Raffaella had that specific thought, and not the utterance time.

According to Sells (1987) and Sigurdsson (1990), morphosyntactic phenomena, such as the *consecutio temporum* and anaphor binding are a result of the viewpoint that the utterance encodes. The notion of viewpoint is to be understood as “the source of the report, the person with respect to whose consciousness (or “self”) [...] and [...] from whose point of view the report is made” (Sells, 1987, p. 445).

According to this argument, subordinate clauses selecting the subjunctive often fail to establish a reference to the utterance time because they convey the viewpoint of an internal subject, an antecedent

3 Semantic accounts of the subjunctive mood in the Romance languages

from whose mental state and point of view the utterance is made. This is distinct from the viewpoint of the external speaker or hearer which one finds encoded in the temporal information conveyed by embedded indicatives. The subjunctive has thus been called a *logophoric mood* (Bianchi, 2001, p. 36) because it refers back to an antecedent, the matrix subject.

Notice that this analysis is not dissimilar from the proposal first given in Farkas (1992b) and further developed in Giorgi and Pianesi (1997) according to which the subjunctive anchors the embedded utterance to the world of a subject, $W_{R(Su)}$, as opposed to the world of the speaker itself, $W_{R(Sp)}$.

This internal perspective is also claimed to be encoded in non-finite clauses. It is no coincidence that non-finite utterances also lack the ability to convey temporal information independently of the matrix clause. Crucially, embedded infinitive and embedded subjunctive utterances share additional similarities concerning co-reference and anaphor binding. Bianchi (2001, p. 36) notes that they both allow long-distance binding: the anaphor in the embedded utterance can refer back to a distant antecedent in the matrix clause, the matrix subject. She gives examples in Italian and Icelandic reported here in (73a) and (74a). Long-distance binding of the pronoun to the matrix subject is not possible with the indicative (see (73b) and (74b)). Similar examples are given in Sells (1987, p. 476). Based on these observations, the subjunctive has been considered to exhibit properties of a non-finite mood.

3 Semantic accounts of the subjunctive mood in the Romance languages

(73)

- a. Gianni_i crede [che Maria odi i propri_i genitori].
Gianni believes that Maria hate.SUBJ the self parents.
“Gianni believes that Maria hates his parents”
- b. Gianni_i sa [che Maria odia i propri*_i genitori].
Gianni knows that Maria hates.IND the self parents
“Gianni knows that Maria hates *his parents.”

(74)

- a. Jon_i segir aD Maria elski sig_i.
J. says that Maria love.SUBJ self.
- b. * Jon_i veit aD Maria elskar sig_i.
J. knows that Maria loves.IND self

Departing from the idea of a complete lack of tense specification, scholars have highlighted that the subjunctive also presents possibilities for the construction of a more complex temporal relationship between the two clauses on the one hand and for framing the embedded event in relation to the speaking time on the other.

Temporal anteriority and, in some instances, even posteriority of the embedded event in relation to the main event can be expressed as suggested by Kempchinsky (1986, pp. 72-73), Giorgi and Pianesi (1997), (2004b), Bianchi (2001, p. 35), and Giorgi (2009, pp. 1844-1845) among others. These involve the use of complex verb tenses, like the *passato prossimo* (present perfect) or the *trapassato* (pluperfect), where the auxiliary is marked for tense and the main verb appears in

3 Semantic accounts of the subjunctive mood in the Romance languages

the past participle. Examples given by Bianchi (2001, p. 35) are reported here in (75) and (76). In (75a) the embedded event is interpreted to have taken place before the main event whereas in (76) it is has still not occurred at the time of speech.

(75)

a. Gianni crede/si rammarica [che ieri alle cinque Maria fosse già partita].

“Gianni believes/regrets (PRS IND) that yesterday at five Mary had (IMPF SUBJ) already left”

b. Gianni crede/si rammarica [che Maria sia già partita].

“Gianni believes/regrets (PRS IND) that Mary has (PRS SUBJ) already left”

(76) Ieri Gianni ha ordinato che l'edificio venga abbattuto entro due giorni.

“Yesterday Gianni ordered (PP IND) the building to be (PRS SUBJ) destroyed within two days.”

Giorgi (2009, p. 1842) points out that in such cases the morphology expressing anteriority is marked on the auxiliary verb rather than the main verb. According to her argument, it derives from aspectual properties like perfectivity rather than temporal ones. For a further explanation read note sixteen in Giorgi (2009, p. 1842).

Per contra, it must be noted that the occurrence of a past subjunctive as a complement to a present epistemic verb is not uncommon and by

3 Semantic accounts of the subjunctive mood in the Romance languages

no means ungrammatical, especially if the embedded verb denotes a state, a durative event or a habitual activity. Sentences (77), (78), (79), (80), (81) taken from the itWac and the C-ORAL-ROM corpus show this syntactic structure: the imperfect subjunctive morphology is marked on the main verb of the embedded clause, rather than on an auxiliary and conveys a temporal interpretation according to which the embedded event takes place in a time considered past in relation to the time of the main clause and, by extension, the speaking time. DAR is enforced in such instances.

(77) Io credo che su queste ci potessimo lavorare

“I believe (PRS IND) that we could (IMPF SUBJ) work on those.”
(imedts01, C-ORAL-ROM)

(78) Faceva un lavoro umile, credo che facesse il manovale

“He had a similar job, I believe (PRS IND) he was (IMPF SUBJ) a manual laborer.” (provincia.venezia.it, itWac)

(79) Credo che facesse molto piacere allora a Franco anche perché era uno studioso

“I believe (PRS IND) that it made (IMPF SUBJ) Franco really happy then also because he was a scholar.” (psychiatryonline.it, itWac)

(80) Era un ex prete e pare che parlasse meglio degli altri.

3 *Semantic accounts of the subjunctive mood in the Romance languages*

“He was an ex-priest and it seems (PRS IND) that he talked (IMPF SUBJ) better than the others.” (parlamento.it, itWac)

(81) Mi pare che frequentasse un’università in Svizzera e che suo padre si accollasse le spese.

“I believe (PRS IND) that he attended (IMPF SUBJ) a university in Switzerland and that his father paid (IMPF SUBJ) all the expenses.” (literary.it, itWac)

Additionally, Giorgi (2009, p. 1844) notes that in complements to factive/emotive verbs the subjunctive can freely select temporal adverbials encoding the past, present and future. As an example, she gives sentence (82).

(82) A Gianni dispiaceva che Maria partisse ieri/ oggi/ domani.

“Gianni was (IMPF IND) sorry that Maria left (IMPF SUBJ) yesterday/ today/ tomorrow.

Yet, sentence (82) was actually perceived to be ungrammatical or at least odd with the adverbs *today* and *tomorrow* referring to the present and future by at least half the native speakers of Italian interviewed for the current analysis. Judgments remain mixed. Nonetheless, the same speakers found the subjunctive past to be much more acceptable with present and future time modifiers when it occurred in complements to epistemic verbs like in (70) presented above. The factive/emotive class

3 Semantic accounts of the subjunctive mood in the Romance languages

again appears to behave slightly differently from the other verb classes that require the subjunctive.

From the evidence gathered, the subjunctive cannot be accurately portrayed as a system completely devoid of the ability to encode temporal meaning. There are possibilities for constructing a more complex temporal relationship between the matrix and the embedded clause selecting a subjunctive. However, they appear to be more limited than those available in indicative embedded utterances, especially in certain clauses like complements to desiderative verbs (*want, hope, wish*) (Picallo, 1984; Bianchi, 2001, p. 35; Kempchinsky, 2009, pp. 1800-1801) where a more fixed temporal syntax is enforced. This is the case for Italian, but it cannot be generalized to other Romance languages. Spanish and Catalan, for one, do not exhibit the same temporal constraints in desiderative subjunctives (Quer, 1998, p. 34; Kempchinsky, 1986, p. 69). Thus, in certain embedded structures DAR is possible with the subjunctive. In this respect, the mood acts like finite moods.

Based on these observations, Bianchi (2001) concludes that “the subjunctive partakes of the nature of both finite and infinitival moods” (p. 36) in the following way: as a finite mood, the subjunctive is anchored to an external perspective, the *external logophoric center*. This is the perspective of the speaker. In such cases, the time reference of the embedded clause is free and *consecutio temporum*, the tense harmony between the two clauses, is not obligatory (see (75), (76)). In other syntactic environments, however, in addition to the perspective of

the speaker, the mood, similar to non-finite moods, can encode an internal perspective, the one of the matrix subject, whose temporal reference and cognitive state it reflects. This happens in the so-called lexical subjunctives, i.e. subjunctives in complement clauses governed by matrix verbs that adhere to the constraint of the consecution *temporum* and allow for long-distance anaphor binding. In Bianchi's terms, these clauses are anchored to both an *internal logophoric center*, i.e. the perspective of the matrix subject, and an *external logophoric center*, i.e. the perspective of the speaker (sentence (72)). The notion of logophoric center has been developed to formalize the intuition that the viewpoint encoded in the sentence is responsible for a whole host of syntactic phenomena like *consecutio temporum* and long-distance anaphor binding. Further, according to Bianchi's argument, the complex syntactic structure of subjunctive clauses which combine an external and an internal logophoric center is responsible for another peculiar syntactic phenomenon concerning anaphor binding, namely the subjunctive disjoint reference constraint.

3.2.2 Subjunctive disjoint reference constraint

Certain subjunctive clauses do not allow coreference between the matrix subject and the subject of the embedded clause. This constraint is known as disjoint reference. Sentences (83) and (84) show how this syntactic effect constrains the interpretation of the verb person of the embedded subjunctive.

3 Semantic accounts of the subjunctive mood in the Romance languages

(83) [I_{o_x}] spero che *pro**_{x/j} porti

[I] hope that PRO bring.*1SG.PRS./3SG.PRS.SUBJ

tutti i documenti con sé.

all the documents with self.

“I hope that she/he brings all the documents with her/him.”

(84) Luca_j spera che *pro* faccia_{x/*j}

Luca hopes that pro do.1.SG.PRS./*3SG.PRS.SUBJ

tutto da sola.

everything alone

“Luca hopes that I do everything alone”

In order to explain the disjoint reference constraint, Bianchi (2001) draws again on the peculiar logophoric nature of subjunctive-selecting argument clauses. These have been shown to combine characteristics of both finite and non-finite clauses due to the fact they encode both an internal perspective, called internal logophoric center, and an external perspective, an external logophoric center.

In addition to the contrasts concerning anaphor binding discussed in the previous section, finite and non-finite subordinate clauses crucially differ in the constraints they place on argument reference. Finite clauses allow for a referentially independent DP or pronoun in subject position: the subject of the embedded clause does not necessarily need to refer

back to an entity introduced in the matrix clause. The option available to the finite clause of selecting a referentially independent argument, i.e. an external participant, is captured by Bianchi with a [+R] feature. This feature, according to her, is directly licensed by the external logophoric center encoded in finite utterances. In sentence (85), the argument in subject position in the subordinate finite clause encodes a referent *nessuno* (“nobody”) that has not been introduced in the matrix clause. It is an external participant.

(85) Ho deciso che più nessuno avrebbe preso il tuo posto.

“I decided that nobody would have taken your spot.” (agi.it itWac)

In embedded infinitives, on the other hand, the subject of the matrix clause controls the embedded sentence (86).

(86) Ho deciso di andare.

“I decided to go.”

In the rare case that the subject is realized in the embedded utterance, this can’t be referentially independent. It has to refer back to the matrix subject. Bianchi (2001, p. 16) shows this constraint of coreference in sentences (87 a, b, and c) and formally captures it by endowing infinitives with a [-R] feature licensed by the internal logophoric center. The [-R] feature captures the constraint that only internal participants can function as the subject of the subordinate infinitive clause.

3 Semantic accounts of the subjunctive mood in the Romance languages

(87)

- a. Ho deciso [di andarci io].
have.1sg decided to go-there I.nom
“I decided to go there myself.”
- b. *Ho deciso [di andarci tu].
*have.1SG decided to go-there you.NOM
- c. *Ho deciso [di andarci Maria].
*have.1.SG decided to go-there Maria

According to Bianchi, subjunctives in complement clauses governed by matrix verbs that respect the *consecution temporum* encode both an external perspective (external logophoric center) with reference feature on the subject [+R] and an internal perspective (internal logophoric center) with [-R] reference feature. Bianchi proposes that the [-R] feature of the internal logophoric center which selects only the internal participants as referents, i.e. the matrix subject, is blocked by the [+R] feature of the external logophoric center that selects external participants. Thus, the subject of the embedded clause cannot take as referent the internal participant encoded in the matrix subject. Co-reference between the matrix subject and the subject of the embedded clause is blocked.

3 Semantic accounts of the subjunctive mood in the Romance languages

Bianchi's crucial point is that not every argument clause that selects a subjunctive exhibits this complex logophoric structure. Also, only clauses that encode the internal logophoric center in addition to the external exhibit disjoint reference. Along with this constraint, they will be showing *consecutio temporum* and consequently no double access reading, an additional effect of the presence of the internal logophoric center. She goes on to verify that these two features indeed correlate.

(88)

- a. Gianni_i sospettava [che e_i fosse stato licenziato].

Gianni suspected that (he) had.SUBJ been fired

- b. (?) Gianni_i sospetta [che ieri alle cinque e_i fosse già stato licenziato]
(anche se gli_ie l'hanno detto solo oggi)

Gianni suspects that yesterday at five (he) had.SUBJ already been fired (although they told him about it only this morning) (taken from Bianchi [2001, p. 40])

Analyzing language data like the sentences reported in (88), Bianchi (2001, p. 40) concludes that the disjoint reference constraint positively correlates with *consecutio temporum*. Whenever the subjunctive clause takes reference to the utterance time and there is no tense harmony between the matrix verb and the embedded subjunctive, the disjoint reference is not enforced (see sentence (88b). In argument clauses that exhibit *consecutio temporum* the disjoint reference is strong. This is the case of the complements to desiderative verbs in (83) and (84). Overall, subjunctives licensed by desiderative predicates constitute the

3 *Semantic accounts of the subjunctive mood in the Romance languages*

environments where both the *consecutio temporum* and the disjoint reference constraints are the strongest and most difficult to override (Farkas, 1992a, p. 86; Kempchinsky, 1986; Laka, 1990; Bianchi, 2001). In subjunctive clauses governed by other predicate classes, such as epistemics, both these constraints are less rigid and can be violated ((88a), (75) above).

Farkas (1992a) had already proposed an account for the disjoint reference constraint which did not rely on the notion of logophoric centers. She suggested that the restriction on coreference arises from the competition between the subjunctive and the infinitive. When the subjects of the two clauses refer to the same entity, the infinitive clause is systematically selected because it is the most parsimonious variant in terms of structural economy. Thus, the subjunctive clause never comes to be used in those circumstances. This type of explanation also resorts to blocking. The subjunctive is blocked by the competing infinitive structure. Bianchi (2001) suggests that such an account fits just as well with the postulation of logophoric centers. The embedded subjunctive, encoding both the external and internal perspective has a more complex internal structure than the infinitive. The latter, being endowed with the internal logophoric center only, suffices to convey the internal perspective of the matrix subject and it effectively blocks the selection of the more cumbersome and unnecessary structure. Such an account brings to the forefront the principled variation between the subjunctive and the infinitive in speech production.

3 Semantic accounts of the subjunctive mood in the Romance languages

However, as already shown above, in speech the indicative is very often selected in the argument clause when the subjects of the two clauses co-refer. This does not happen with sentences that encode negative polarity only, i.e. negated predicates (see (59) and (57b) above), but in affirmative contexts as well. Sentences (89), (90), (91) and (92) demonstrate that the indicative does not have the same limitation on coreference that the subjunctive has and speakers resort to it when the subjunctive is no longer an option.

(89) Penso che vado in ferie e sto via per tre settimane

“I think that I will go (IND) on holiday and I will stay there for three weeks” (blog.libero.it, itWac)

(90) Lei pensa che fa bene alla figlia

“She thinks that she helps (IND) her family”(dramma.it, itWac)

(91) Spero che mi sono spiegato.

“I hope that I explained (IND) myself well” (biggame.it, itWac)

(92) [...] per tutti coloro che credono che hanno una sensibilità [...].

“[...] for those who believe (IND) that they have a sensibility[...].”
(provincia.fit.it, itWac)

3 *Semantic accounts of the subjunctive mood in the Romance languages*

The exception is constituted by strong desiderative and directive predicates which exhibit a robust disjoint reference constraint either with the subjunctive or with the indicative (see (93)).

(93) **PRO* voglio che vado a casa prima di pranzo.

*I want that I go home before lunch.

According to the thesis proposed in Bianchi (2001) and others (Sells, 1987; Sigurdsson, 1990), indicative clauses license the external perspective of the speaker only. This is syntactically captured in the so-called external logophoric center.

The internal viewpoint of the matrix subject and the external perspective of the speaker coincide in the first-person *I* or *we*. In those instances, adopting Bianchi's approach, one could argue that the indicative clause with its external logophoric center may be selected because the perspective of the matrix subject and the one of the external speaker are one and the same (see sentences (89) and (91)). Indeed, the indicative may be preferred to the subjunctive as the more economic structure.

Still, the theory fails to explain why speakers tend to choose the indicative in the complement clauses exhibiting coreference between the matrix subject and the subject of the embedded verb when these subjects do not refer to the external speaker. An example is subjects

3 Semantic accounts of the subjunctive mood in the Romance languages

encoding a third-person referent *she, he, they* (sentences (90), (92) above).

The difficulty in modeling how syntactic phenomena like the subjunctive disjoint reference constraint operate has already been noted already in the literature (Landau, 2004). Many additional factors that seem to affect the inhibition of subject coreference have been suggested. Moreover, judgments on coreference vary among speakers and are not categorical: some sentences sound more acceptable than others. They also vary in relation to the language and language variety considered, even within the Romance family (Kempchinsky, 2009). Thus, Ruwet (1991) shows how in French even with the verb *vouloir* “want”, that is commonly associated with a strong disjoint reference constraint, the acceptability of coreference can be modulated, among others, with:

- (a) passive subjects
- (b) modal verbs in the subjunctive clause
- (c) focused pronominal subject in the subjunctive clause

The relevant examples are given in (94) and can also be found also in Kempchinsky (2009, p. 1792).

(94)

- a. * Je veux que je parte.
“*I want that I leave (SUBJ).”
- b. ? Je veux que je sois autorisé à partir demain.
“I want that I be (SUBJ) authorized to leave tomorrow.”

3 Semantic accounts of the subjunctive mood in the Romance languages

c. Je voudrais bien que je puisse enfin être autorisé à partir.

“I would certainly that I should (SUBJ) finally be authorized to leave.”

Kempchinsky (2009) makes similar sentences in Spanish and presents them to native speakers. Their judgments differ somewhat from the scale of preference reported in Ruwet. Most reject a coreferential interpretation of the subjects in all sentences in (95). However, sentence (95a), with a passive subject in the embedded clause is accepted the most.

(95)

a. Ana quiere que sea elegida.

“Ana wants that (she) be (SUBJ) chosen.”

b. Ana quiere que pueda acompañaros.

“Ana wants that (she) be able (SUBJ) to accompany you.”

c. Ana quiere que ELLA os acompañe.

“Ana wants that (she) accompany (SUBJ) you.”

d. Ana quiere que os acompañe.

“Ana wants that (she) accompany (SUBJ) you.”

Similarly, in Italian most speakers reject a coreferential interpretation of the subjects in sentences in (96) with the matrix verb *volere*. Nonetheless, the rejection rate as a measure of the strength of the disjoint reference constraint increases from a. to d.

(96)

3 Semantic accounts of the subjunctive mood in the Romance languages

- a. Carla vuole che sia ascoltata.
“Carla wants that she be (SUBJ) listened to.”
- b. Carla vuole che possa agire senza intromissioni esterne.
“Carla wants that she be (SUBJ) allowed to act without external interferences.”
- c. Carla vuole che lei stessa presenti la squadra.
“Carla wants that she herself introduce (SUBJ) the team.”
- d. Carla vuole che presenti la squadra.
“Carla wants that she introduce (SUBJ) the team.”

Further theories sought to reduce the complexity to an underlying principle in the spirit of the works reviewed in detail above, either within the framework of Binding Theory or by presupposing a functional competition between the infinitive and the subjunctive construction (Piccolo, 1985; Rizzi, 1990; Progovac, 1993, 1994; Schlenker, 2005). Most recently, Costantini (2016) called on the principle of self-knowledge or introspection. According to him, the disjoint reference results from a clash between the embedded clause that encodes direct knowledge tapped within the self, therefore 100% reliable, and matrix verbs that encode indirect or inferred knowledge, i.e. evidential predicates like *pensare* “think”, *parere* “seem”. This theory as he himself admits may apply to epistemic verbs but not to volitional or emotive-factive predicates.

Farkas (1992a) and Quer (1998) proposed that the subject of the matrix clause cannot “control” or “be responsible” for the event encoded in the

3 Semantic accounts of the subjunctive mood in the Romance languages

embedded clause. However, this account fails to explain why certain matrix verbs exhibit a stronger disjoint reference constraint than others. See the difference between *volere* “want” and *sperare* “hope” in (97) with passive subjects. Coreference is by far more accepted with the latter verb in (97b).

(97)

- a. Licia vuole che sia la prima della lista.
“Licia wants that she is (SUBJ) the first on the list.”
- b. Licia spera che sia la prima della lista.
“Licia hopes that she is (SUBJ) the first on the list.”

Kempchinsky (2009) among others already noticed that the lexical identity of the matrix verb influences coreference in a manner that seems idiosyncratic. She concludes that “a purely structurally based binding theory approach fails to capture the full range of facts” (p. 1792)

Starting from the subjunctives that appear as complements to desiderative and directive predicates, she makes a parallel between the subjunctive and the imperative mood. This analogy is originally found in Portner (2004).

From the point of view of semantics, both imperatives and complements to desideratives encode events or states that have not occurred in the real world, W_R . The difference lies in the point of view, i.e. the modal base of the utterance. Similar to the indicative/subjunctive opposition, the subjunctive encodes the point of

view of the matrix subject while the imperative encodes the point of view of external speakers. In Portner's terms the subjunctive conveys possibilities on the To-Do List of the matrix subject, while the imperative encodes possibilities on the To-DoList of the speaker (Kempchinsky, 2009, p. 1794).

Kempchinsky posits that the disjoint reference is triggered by an underlying syntactic structure that shares similarities to the structure of the imperative (see Kempchinsky [2009] for a detailed account on the 'quasi-imperative operator', where it is located in the syntactic representation and how the latter differs from Bianchi's (2001) logophoric centers).

The imperative can select any person other than the first person singular, the person that refers to speaker alone. Similarly, the subjunctive can encode any person other than the matrix subject alone. Thus, both the imperative and the subjunctive exclude the referent whose point of view constitutes the modal base of the sentence.

Kempchinsky (2009) notes that first person plural imperatives are admitted (see (98)). In the same way, the subjunctive admits an overlapping coreference (see (99) taken from Kempchinsky [2009, p. 1795]). This is also possible in Italian (100).

(98) *Andiamoci tutti insieme!*

"Let us go (IMP) all together!"

3 Semantic accounts of the subjunctive mood in the Romance languages

(99)

- a. Quiero que vayamos allí juntos.
“I want that we go (SUBJ) there together.
- b. Queremos que yo salga primero.
“We want that I leave (SUBJ) first.”

(100)

- a. Voglio che Andiamo tutti insieme alla festa
‘I want that we go (SUBJ) to the party together.’
- b. Vogliamo che io parta per primo.
‘We want that I leave (SUBJ) first.’

Based on these compelling parallels, she argues that the subjunctive governed by desiderative and directive predicates can be treated as an embedded imperative. She mentions additional facts in support of the special relation between the two moods that have been observed in pragmatics and in language acquisition studies. For instance, in Spanish the subjunctive and the imperative appear to be functionally equivalent in discourse (see (101)).

3 Semantic accounts of the subjunctive mood in the Romance languages

(101)

Person A: Sube!

Climb.2.SG.IMP-in (e.g. into a car)

“Climb in!”

Person B: Qué dices?

“What are you saying?”

Person A: Que subas!

That climb.2.SG.SUBJ-in

“Climb in!”

(Kempchinsky, 2009, p. 1796)

Moreover, language acquisition studies in Spanish L1 report that “earliest observed uses of subjunctive clauses are in indirect commands (Montrul, 2005) or with, in particular, the verb *querer* ‘to want’ (see Blake, 1983 among others)”

I briefly add to these observations that in classical Latin the subjunctive was used as an imperative for exhortations, commands and for prohibitions together with the negation operator (Calboli, 1969) in unembedded utterances. See (102), (103) in Cicero’s *Orationes* (Cicero, 1985) and (104) in Cicero’s *Philippics* (Manuwald, 2012), all reported in Tantucci (1962, p. 244). All verb persons of the subjunctive were productive and could indeed supplement the imperative conjugation. The latter only had two verb persons, the second singular and plural (see Table 3.1).

3 Semantic accounts of the subjunctive mood in the Romance languages

- (102) Amemus patriam.
love.1.PL.PRS.SUBJ country.ACC
“Let us love our country.” (Cicero, *Orationes*, *Pro Sestio*, in Cicero [1895, p. 70] reported in Tantucci [1962, p. 244])
- (103) Ne attingant rem publicam.
NEG touch.3.PL.PRS.SUBJ res.ACC publica.ACC
“They must not touch the *res publica*.” (with imperative value: *Do not touch the res publica*.)
(*Orationes*, *Pro Sestio*, in Cicero [1856b, p. 569] reported in Tantucci [1962, p. 244])
- (104) Arma deponat
weapon.ACC.PL lay.3.SG.PRS.SUBJ
“He must lay down the weapons!” (with imperative value: *Lay down the weapons*.)
(Cicero, *Philippics* 3-9, Volume 1, in Manuwald, 2012, p. 200 reported in Tantucci, 1962, p. 244)

3 Semantic accounts of the subjunctive mood in the Romance languages

Table 3.1

Subjunctive and imperative present active conjugation Latin verb amāre “to love”

Subjunctive	Imperative
<i>am-em</i>	
<i>am-es</i>	<i>am-a</i>
<i>am-et</i>	
<i>ae-ēmus</i>	
<i>am-ētis</i>	<i>am-āte</i>
<i>am-ent</i>	

This function is generally no longer available in the Romance languages. However, similar to Latin, in Italian the third-person subjunctive singular and plural forms of the present tense fill the morphological gaps in the imperative present paradigm. They are used for the polite address of courtesy *lei* “You. 3SG” and *loro* “You. 3PL” (see (106) and (107)).

(105) Siediti!

Sit.2.SG.PRS.IMP

“Sit down!”

(106) La prego, si sieda.

PRO please, REFL.PRO sit.3.SG.PRS.SUBJ

“Please, take a seat!” (formal)

3 Semantic accounts of the subjunctive mood in the Romance languages

- (107) Mi ascolto gentilmente!
 pro.ACC listen.3.SG.PRS.SUBJ please!
 “Listen to me please!” (formal)

In the forms presented above, the subjunctive has the illocutionary value of a command and remains as a vestige of the jussive subjunctive in Latin. In this domain, the imperative and the subjunctive thus concur to encode the very same function.

The partial similarities between the syntax of the subjunctive and that of the imperative are taken as a clue to the disjoint reference constraint and to why this effect is particularly strong in desiderative and directive contexts. However, this account cannot address the very distinctive correlation between the syntactic constraint on coreference and each matrix verb, nor can it account for why the disjoint reference would operate in other contexts that are not as closely related to the semantics of the imperative, for instance in subjunctives governed by epistemic or factive/emotive verbs. Kempchinsky limits herself to the very general claim that the subjunctive under volitive verbs should be understood as the core subjunctive. One can deduce that the syntactic properties of this core subjunctive could be passed down somehow to other types of subjunctives in complement clauses.

Finally, the account does not address the relation between the disjoint reference and the subjunctive/indicative alternation established through corpus analyses and presented above.

3.2.3 Complementizer deletion

A well-studied syntactic property related to the subjunctive in Italian is the possibility to omit the complementizer “that”, which connects the main clause to the embedded argument clause. Complementizer deletion (CD) is not possible when the embedded clause selects the indicative (Giorgi & Pianesi, 1997, 2004a; Poletto, 2000, 2001). An example is given in Giorgi (2009, p. 1841) and is reported here in (108). (108)

- a. Gianni ha detto *(che) ha telefonato Maria.
“Gianni said that has (IND) called Maria.”
- b. Gianni credeva (che) avesse telefonato Maria.
“Gianni believed (that) had (SUBJ) called Maria.”
(Giorgi, 2009, p. 1841)

According to Giorgi and Pianesi (2004a) and Giorgi (2009), this behavior of the complementizer sets Italian apart from other Romance languages where CD is not found. However, studies report that the complementizer is often omitted in Quebec French as well (Poplack et al., 2013; Martineau, 1985; Dion, 2003; Warren, 1994). Nonetheless, an important difference between the two languages must be noted. In Quebec French, the complementizer may also be left out also when the indicative occurs in lieu of a normative subjunctive in speech. Sentence (109) taken from Poplack et al. (2013, p. 174) shows this phenomenon.

3 *Semantic accounts of the subjunctive mood in the Romance languages*

(109) Fallait [] tu mets un chapeau pour aller à église.

“It was necessary [] you put (IND) on a hat to go to church.”

(20C.064.2119)

In fact, Poplack et al. (2013) note that, in Québécois, from the beginning of the 20th century, CD positively correlates with the indicative. On the other hand, the probability of selecting a subjunctive form increases when the complementizer *que* is realized.

In Italian, natural speech is also characterized by the indicative/subjunctive variability. The indicative is very often selected in subordinate clauses governed by verbs that, in keeping with their lexical semantics, would require a subjunctive according to standard grammars. However, when the indicative present or past occurs in non-standard speech, the complementizer cannot be omitted. CD is usually perceived as ungrammatical. This is shown in (110) taken from Giorgi and Pianesi (2004a, p. 191).

(110) Gianni credeva *(che) aveva telefonato.

“Gianni believed *(that) he had(IND) called.”

The analysis of the mood variability in the speech corpus C-ORAL-ROM (Cresti & Moneglia, 2005) confirms that the complementizer is always realized when the embedded verb is marked for the indicative past and present. On the other hand, CD is not unusual when an

3 Semantic accounts of the subjunctive mood in the Romance languages

epistemic or an optative verb selects an embedded future in speech (see (111) and (112)).

(111) Ad aprile non credo troverai molto da fare. (nonlosopappe.it, itWac)

“In April I don’t think you will find (FUT) much to do.”

(112) [...] ma spero avrà una grande cucina funzionale e attrezzata.

“but I hope it will have a big, efficient kitchen.”

(spaziocucina.surjaring.it, itWac)

CD occurs significantly more often if the matrix clause is itself a subordinate clause, independently of the lexical identity of the matrix verb and of the embedded verb. There is a strong tendency to avoid the repetition of the subordinating conjunction that, to the best of my knowledge, has not been discussed so far in the literature. This effect can be established both with the embedded indicative future ((113), (114), (115)) and the embedded subjunctive (examples (116), (117), (118)).

(113) Procederò a una rilettura che spero avrà esiti più felici.

“I will proceed to a second reading which I hope will have (FUT) better outcomes.” (centomovimenti.it, itWac)

3 Semantic accounts of the subjunctive mood in the Romance languages

(114) È talmente evidente la lacuna che penso sarà al più presto colmata
“The deficiency is so evident that I think it will soon be addressed
(FUT).” (fiorentina.it, itWac)

(115) Peccato per la telecronaca italiana che immagino sarà penosa
quest’anno
“Shame for the Italian television news which I imagine will be
(FUT) lousy this year” (dvd.it, itWac)

(116) Non partecipa alla corsa che speravo facesse.
“He does not take part in the race which I hope he would do
(SUBJ).” (cicloweb.it, itWac)

(117) Mi ci è voluto molto per capire che la persona che credevo fosse
lui in realtà non esisteva.
“It took a long time for me to understand that the person that I
thought he was (SUBJ) in reality didn’t exist.” (girlpower.it,
itWac)

(118) Erano stati affissi manifesti del governo che si pensava fossero
utili a calmare la folla
“During the night a lot of posters were put up from the
government that they thought would be (SUBJ) useful to calm the
crowd.” (cronologia.it, itWac)

3 Semantic accounts of the subjunctive mood in the Romance languages

In violation of the strong constraint disallowing CD with the indicative present and past, in (119) we also find the unusual omission of the complementizer with the indicative past, one of only few instances in the 1.5-billion-word corpus ItWac. In all such cases, the matrix clause is in fact a subordinate clause already introduced by the complementizer *che*. The disinclination toward realizing the complementizer twice appears to be so high as to supersede other rules concerning the occurrence of the complementizer.

(119) Il partito comunista di Russia che credo aveva – non ricordo – circa 15 o 20 delegati.

The communist party in Russia which I believe had (IND) – I don't remember – roughly 15 or 20 delegates.

Such a phenomenon already suggests that the deletion of the complementizer does not depend on the presence of the subjunctive alone. Giorgi and Pianesi (2004a) and Giorgi (2009) unearth additional semantic and syntactic properties of the embedded clause that correlate with CD:

- (a) Factive/emotive verbs do not allow complement deletion, an additional characteristic that sets this class apart from other classes that select the subjunctive.
- (b) There is a positive correlation between complement deletion and the *consecutio temporum* between the matrix verb and the

3 Semantic accounts of the subjunctive mood in the Romance languages

subjunctive (Giorgi, 2009, p. 1841; Giorgi and Pianesi, 2004a, 2004b).

Putting together (b) with the correlation between *consecutio temporum* and the disjoint reference constraint established in the previous section, we may propose an additional property concerning the syntax of subjunctive argument clauses:

- (c) There is also a correlation between CD and the disjoint reference constraint.

In the following, I review and discuss properties (b) and (c) relevant for the present discussion.

Giorgi (2009, p. 1845) shows through a number of linguistic examples reported here in (120) that complement deletion is impossible in the subjunctive clauses that violate *consecutio temporum*. This is the case of (120c), which features a present subjunctive governed by a past indicative, the only utterance where the omission of the complementizer would result in ungrammaticality. Recall that the violation of *consecutio temporum* is also associated with the weakening or canceling of the subjunctive disjoint reference, feature (c). In (121), which slightly modifies (120) to encode a female subject, an interpretation in which the matrix subject and the subject of the embedded clause are coreferent is possible: *Gianna hypothesized that*

3 Semantic accounts of the subjunctive mood in the Romance languages

she herself may be pregnant. Here as well, *consecutio temporum* is violated and complementizer deletion is impossible.

(120)

- a. Gianni ha ipotizzato (che) fosse incinta.
“Gianni hypothesized (that) (she) was (PST SUBJ) pregnant.”
- b. Gianni ipotizza (che) sia incinta.
“Gianni hypothesizes (that) (she) is (PRS SUBJ) pregnant.”
- c. Gianni ha ipotizzato *(che) sia incinta.
“Gianni hypothesized (that) she is (PRS SUBJ) pregnant.”
(Giorgi, 2009, p. 1845)

(121) Gianna ha ipotizzato *(che) sia incinta.

“Gianna hypothesized that she is (PRS SUBJ) pregnant”

Giorgi and Pianesi (2004a) find that the crucial property driving the correlation between *consecutio temporum* and the complementizer deletion is the double access reading itself. When the DAR is enforced, i.e. when the temporal coordinate of the external speaker is present and the embedded clause is evaluated both in relation to the time of the main event and to the utterance time, the complementizer cannot be omitted.

In order to explain why the DAR inhibits complementizer deletion, the two scholars posit that the complementizer has two distinct functions in DAR clauses (e.g. embedded indicatives) and in non-DAR clauses (i.e. embedded subjunctives that observe the *consecutio temporum*). They are actually two functionally distinct

3 Semantic accounts of the subjunctive mood in the Romance languages

complementizers. To back up this claim, they remark that in many languages the complementizer that selects the indicative and the one that selects the subjunctive are also lexically distinct. This is the case in Greek (Roussou, 2000) but also in Romanian, a romance language (Farkas, 1984; Dobrovie-Sorin, 2011; Kempchinsky, 2009).

In subjunctive sentences with no DAR, the complementizer is part of the subjunctive morphology:

[...] the Italian subjunctive exhibits a sort of discontinuous morphology, including both the verbal ending and the complementizer. The two can either be realized together (i.e., syncretically, adopting Giorgi and Pianesi's terminology) or scattered, in which case the word *che* appears in the embedded clause.

The complementizer thus only encodes the feature MOOD. Deleting it would not result in any loss of information. The MOOD feature would then be conveyed by the verb's inflectional ending together with the feature TENSE. This is shown in the syntactic representations in (122) and (123) taken from (Giorgi, 2009, pp. 1847-1848). They respectively depict a scattered realization and a syncretic realization of mood. In the latter, the complementizer is not realized.

(122) Gianni credeva che Maria dormisse.

“Gianni believed that Maria slept (PST SUBJ)”

[. . . . [V credeva [_{MOODP {+mood}} che [_{TP . . . dormisse_{+past} . . .]]]]}

3 Semantic accounts of the subjunctive mood in the Romance languages

(from Giorgi, 2009, p. 1847)

(123) Gianni credeva dormisse.

“Gianni believed she slept (PST SUBJ).”

[. . . . [v credeva [MOOD/TP dormisse_{+mood; +past} . . .]]]

(from Giorgi, 2009, p. 1848)

The complementizer introducing indicative argument clauses fulfills an entirely different function: it encodes the speaker’s temporal coordinate. This feature is represented with the symbol Σ in the syntactic structure provided by Giorgi and given in (124).

(124) Gianni ha detto che Maria ha telefonato.

“Gianni said that Maria has (IND) called.”

[. . . . [v detto [CP . . . che Σ [TP . . . T σ . . . [. . . ha telefonato_{ Σ ; σ } . . .]]]]]

(Giorgi, 2009, p. 1848)

In order to instantiate DAR, both the temporal coordinate of the internal matrix subject (feature σ in the structure in (124)) and the speaker’s temporal coordinate are necessary. Thus, the deletion of the complementizer is not possible.

The subjunctive clauses that violate *consecutio temporum* and instantiate double access reading, similar to indicative clauses, rely on the complementizer to encode the speaker’s temporal coordinate. For

3 Semantic accounts of the subjunctive mood in the Romance languages

this reason, the complementizer cannot be deleted. Their structure is given in (125).

(125) Gianni ha ipotizzato che sia incinta.

“Gianni hypothesized that she is (PRS SUBJ) pregnant.”

[. . . [V hypothesized [_{CP} che _Σ [_{MOODP} sia_{+mood; +pres} . . .]]]]

(Giorgi, 2009, p. 1849)

Giorgi and Pianesi argue that “the DAR is actually enforced in all cases where an indicative appears [and] it is empirically detectable with the present tense and the future [indicative]” (Giorgi, 2009, p. 1840).

Focusing on the future tense, as already discussed, the embedded future yields an interpretation according to which the embedded event has not taken place either at the time of the main event or at the utterance time. Giorgi (2009, p. 1840) shows this in sentence (126): there is no possibility that the buying of the house has already occurred at the time of the utterance. DAR must be enforced.

(126) Gianni ha detto *(che) Maria comprerà casa

“Gianni said that Maria will buy (FUT) a house.”

The future is thus a case in point where the double access reading is enforced. Giorgi and Pianesi’s syntactic representation makes the complementizer essential for DAR because it is invested with the role of encoding the speaker’s temporal coordinate. Against this theory,

natural language data shows that complementizer deletion is actually possible with certain matrix verbs that select future morphology. Examples of CD with the future tense governed by optative or epistemic verbs were already given in (111) and (112). In (127) I provide another instance. The omission of the complementizer is not perceived to be ungrammatical in such sentences.

(127) Credo faremo seguire un corso per le emergenze ai nostri docenti.

“I think we will have teachers follow a course for emergencies.”

(centroprociv.it, itWac)

Much like in the case of Bianchi’s theory, it is possible to argue for the special status of first-person subjects in a syntactic theory that relies on the distinction between internal subject and external speaker encoded in the modal base. Indeed, in a sentence like (127) with a first-person matrix subject, the temporal coordinate of the main event and the speaking time must coincide: The thought of having a course and the communication thereof are simultaneous.

The temporal coordinate of the internal subject, feature σ in the structure proposed in (128), thus suffices to generate the double access reading: the embedded event will be evaluated and considered future in relation to the speaking time as well. Such an explanation however entails that also the indicative can encode an internal perspective which ultimately allows the complementizer to be deleted.

3 *Semantic accounts of the subjunctive mood in the Romance languages*

(128) [. . . . [V *credo* [TP . . . T_σ . . . [. . . *faremo*_{σ} . . .]]]]

The data coming from the itWac corpus suggests that the first person does indeed have a special status in relation to the complementizer. Considerably more frequent are the complementizer deletions in argument clauses governed by optative or epistemic verbs which select the indicative future or even present (129) when the matrix subject encodes a first-person referent *I* or *we*.

(129) *La faccenda credo ha un interesse oltre quello strettamente tecnico e professionale.*

“I believe the matter has an interest that goes beyond the strictly technical and professional.” (girodivite.it, itWac)

However, complementizer deletion with future tense can also be found when the matrix subject encodes other referents (see (130)). These uses may be more infrequent or in some instances more marked. Nonetheless, they cannot be straightforwardly reconciled with the rationale proposed in Giorgi and Pianesi (2004a) and further expanded in Giorgi (2009).

(130) *Le femmine non credono sarà loro di nessuna utilità*

“Women do not believe it will have any benefit.” (scienzaesperienza.it, itWac)

3.2.4 Conclusion

Bianchi's (2001) work focused on the relation between the *consecutio temporum* and the disjoint reference constraint whereas Giorgi and Pianesi (2004a) addressed the relation between *consecutio temporum* and the deletion of the complementizer (CD) in subjunctive complement clauses.

Both studies draw on the double access reading property and the distinction between the viewpoint of the speaker and the viewpoint of the internal subject to explain each individual relation. Moreover, both studies adopt the same theoretic framework, the minimalist program (Chomsky, 1995, 1998, 2008), to develop their account.

Nonetheless, a unifying theory that tackles the relation between all three variables is not attempted even though all three morphosyntactic features – *consecutio temporum*, subjunctive disjoint reference constraint, complementizer deletion – actually correlate and the same principles are called upon to understand each relation pairwise.

Moreover, Bianchi's account cannot be reconciled a posteriori with Giorgi and Pianesi's theoretical framework. Table 3.2 shows how the two theories conceptualize the syntactic structure of the subjunctive clauses that obey the *consecutio temporum* and of those that violate it. Bianchi posits that the clauses that obey the *consecutio temporum* have a complex structure, encoding both the viewpoint of the external speaker and the viewpoint of the internal matrix subject. On the contrary, according to Giorgi and Pianesi, this dual perspective is encoded in the clauses that violate the tense harmony.

3 Semantic accounts of the subjunctive mood in the Romance languages

Table 3.2

The subjunctive explained by generative theories

	subjunctive clause violating <i>consecutio temporum</i>	subjunctive clauses obeying <i>consecutio temporum</i>
Bianchi (2001)	viewpoint of speaker	viewpoint of the matrix subject and viewpoint of speaker ⁵
Giorgi and Pianesi (2004a) Giorgi (2009)	viewpoint of the matrix subject and viewpoint of speaker ⁶	viewpoint of the matrix subject

Finally, the works at the interface between syntax and semantics closely reviewed above fail to give a complete picture of the factors affecting the mood distribution and the syntax of subjunctive and indicative clauses in natural speech. Nevertheless, these studies have the benefit of highlighting fundamental characteristics of subjunctive clauses. In particular, they brought to light and examined a non-trivial

⁵ Bianchi (2001, pp. 11-13) in her syntactic representation calls the two viewpoints internal logophoric center and external logophoric center.

⁶ These two viewpoints are captured in the syntactic representation suggested by Giorgi (2009, pp. 1848-1849) by feature σ , which represents the temporal coordinate of the matrix subject and by feature Σ , which encodes the temporal coordinate of the external speaker.

3 Semantic accounts of the subjunctive mood in the Romance languages

correlation between complementizer deletion, *consecutio temporum* and the disjoint reference constraint.

Moreover, Kempchinsky (2009), based on Portner (2004), highlights the partial similarities between the subjunctive and the imperative mood in both syntax and semantics and ultimately makes a case for a core subjunctive, one that is learnt first and best retained in the face of attrition, namely the subjunctive that appears in complements to desiderative and directive predicates.

Having reviewed the main theoretical accounts of the Romance subjunctive, in the next chapter I provide an overview of the syntax and functions of the subjunctive in Latin along with clues concerning the diachronic changes that must have affected the mood from Vulgar Latin to the Romance languages. The semantic theories presented so far will be compared to the accounts of the subjunctive mood written by the ancient grammarians in order to highlight to what extent the effort to establish a core underlying meaning of the mood comes from the grammatical tradition and whether it captures the subjunctive/indicative distribution in Latin.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

In the previous chapter, I explored a number of refined and often complex theories at the interface between semantics and syntax that have been put forwards in order to explain the equally complex and seemingly quirky properties of the subjunctive mood in the Romance languages. Although such theories differ in some crucial aspects concerning the syntactic representation that licenses the subjunctive mood, most highlight the role of the modal base or viewpoint. The modal base is considered to represent either the source, i.e. the individual, based on whose knowledge the proposition is evaluated, or more generally, the set of propositions against which the utterance is appraised.

The subjunctive is oftentimes interpreted as signaling a shift in this modal base from the default, characterized by the knowledge and viewpoint of the speaker and addressee, whose set of assumptions frames the discourse, to an internal source, the matrix subject, whose viewpoint frames the embedded utterance. Other times, the subjunctive is understood to shift the modal base from the actual world, the set of propositions that apply in $W_{(R)}$, to a set of possible worlds, $W_{(P)}$, against which the utterance is evaluated.

Both accounts have been most popular for understanding the occurrence of the subjunctive in subordination, mainly in argument clauses. Scholars have recognized that such models cannot be applied to all the instances in which the subjunctive actually occurs in speech.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Some works have ventured the idea that different types of subjunctives exist. These follow different rules and signal different things (recall the various labels given: for instance, polarity subjunctive, lexical subjunctive etc.). They acknowledge that the taxonomies or analyses proposed so far remain partial. Nonetheless, most studies more or less tacitly assume that a core subjunctive semantics or a core subjunctive property can be identified. Others suggest that there are some core meanings of the subjunctive and more peripheral ones. For instance, the intentional subjunctive, which appears under volitive or directive predicates has been proposed as the core subjunctive, one that, at least in theory, resists crosslinguistic variation as it appears in all Romance languages.

The focus on the semantic interpretation(s) of the subjunctive/indicative opposition present in the works in theoretical linguistics is emphasized in the accounts provided in prescriptive grammars and grammar teaching books employed in schools (Poplack, 1990; Poplack et al., 2013). As was already shown in the previous chapter, these often borrow the functional categories used to describe the subjunctive in Latin and present a detailed list of subjunctive types which are mainly based on the lexical semantics of the governor verbs or the subordinating conjunctions. Some of those categories are reported again in the following:

- 1) Epistemic or dubitative subjunctive
- 2) Optative or volitive subjunctive
- 3) Jussive or directive subjunctive

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

- 4) Factive/emotive subjunctive
- 5) Concessive subjunctive

Such efforts of abstraction involving the notation of ever more detailed semantic values expressed by the subjunctive morphology have also been noted by Poplack (1990) and Poplack et al. (2013) while reviewing the grammars that describe or codify standard French over the centuries. They cite among others Bidois & Bidois (1971), Denis & Sancier-Chateau (2004) and Leeman-Bouix (1994). For a more extensive review of how the grammars of French treat the subjunctive see Poplack et al. (2013, pp. 150-155). They all share the fundamental conceptualization of the subjunctive as a mood functionally opposed to the indicative in that it marks the realm of possibility and subjective representation in contrast to a factual and a referentially-free representation.

Against this backdrop, more data-driven research, including my own analyses, shows how the alternation between the subjunctive and other moods, like the indicative and the conditional, cannot be properly grasped by relying solely on semantic-based accounts. Mood variation across languages and within the idiom of a single community or a single speaker suggests that additional forces shape speakers' mental representation and use of the subjunctive. These hidden factors remain unexplored by the studies presented so far.

The analyses of the development of the subjunctive across time reveal that the system has undergone a series of changes and

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

simplifications. They have affected its use and the mood distribution in all Romance languages (Harris, 1974; Bergareche, 1990; Poplack, 1990; Silva-Corvalán, 1994b; Bybee, 2002b). In addition to the structural reorganization, the diachronic studies in French, Spanish and Italian suggest that the subjunctive has been subject to *desemanticization*. This process indicates that the mood is progressively losing or has for the most part already lost the ability to encode the set of functions that used to make up its semantic domain.

The subjunctive is thus viewed as an empty syntactic marker of subordination, devoid of any specific meaning. Indeed, the presence of the subjunctive in unembedded clauses is very scanty and for the most part crystallized in a few formulaic constructions. In parallel, in subordination, there is an overall decline in use. The latter is no longer systematic: “previously obligatory contexts for the use of Sub are now categorically Ind or allow both Sub and Ind to different degrees across social and geographic parameters.” (Silva-Corvalán, 1994b, p. 256). Alternation with other moods is also possible (Klein, 1975; Silva-Corvalán, 1985). These are often younger grammatical strategies that appear as structural innovations in the course of linguistic change (see the conditional mood or the indicative future). Crucially, the mood opposition cannot be motivated on account of the semantic and syntactic environment. The studies on Italian and French point out the methodological flaws in the approach supported by most grammarians who nowadays attribute to the subjunctive the semantic content conveyed by the linguistic elements which co-occur with it in language

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

practice (Blücher, 2003, pp. 169-170 and Poplack, 1990). A similar position has been expressed by Porto Dapena (1991) for the analysis of the Spanish *subjuntivo*, whose modal value is deemed abstract and varied to the extent that carving out a specific semantic domain for this mood results in an arbitrary and inaccurate generalization (21). Research suggests that the loss of the subjunctive is accelerated in situations of language contact (Montrul, 2007; Lynch, 1999; Silva-Corvalán, 1994b) and the mood disappears first in the contexts where the subjunctive/indicative opposition is meaningful. Alongside the gradual loss of meaning and the decline in use, morphological simplification and erosion of phonetic material takes place.

Before turning to the studies that have looked at the mood alternation in contexts that, according to prescriptive grammars, are reserved to the subjunctive, I will give an overview of the functions and domains of application of the subjunctive in Latin, the common ancestor to the Romance languages.

The historical perspective will provide indirect clues to the development that the subjunctive has undergone throughout the centuries in relation to other morphological systems and an important frame of reference in order to understand its role nowadays in the Romance languages. Additionally, I will discuss the metalinguistic reflection of the ancient Roman writers and grammarians concerning the subjunctive in classical Latin and compare it to the most recent theoretical work in order to supply a relevant context for the theories

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

that have attempted to explain the mood opposition in more recent times.

4.1 The notion of the subjunctive among ancient grammarians

Seeing how the conceptualization of the subjunctive promoted in grammar books and in linguistic essays heavily relies on the semantic opposition between the subjunctive and the indicative, it might come as a surprise how different an understanding of this mood was cultivated among the ancient grammarians of Latin and Greek, especially if one takes into account the alleged desemanticization that has affected the subjunctive over the course of its structural evolution from vulgar Latin to the Romance languages of today. They present the subjunctive as the mood of the hypotaxis. It indicates the syntactic position within the sentence, more specifically, the fact that we find ourselves in a subordinate clause introduced by a conjunction. This is its defining feature epitomized in the Greek name itself: ὑποτακτική (*hypotaktikḗ*) meaning “arranging underneath” (see Basile [2001, pp. 431-442] and Calboli [1969] for a comprehensive and detailed account on the history and the syntax of the subjunctive mood in Latin and Greek. They provide the historical background in the current analysis). The Latin term *subiunctivus* or *coniunctivus* loaned from Greek highlights this property as well. According to Macrobius (GL V. 643. 22), it means “ex sola coniunctione, quae ei accidit”, that is, “the mood of the conjunction it goes with”. Such a definition takes reference to the

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

syntagmatic relation between the subjunctive and another element of the sentence, the conjunction. Any semantic value of the mood is explicitly rejected. Thus, Diomedes Grammaticus, in his *Ars Grammatica*, defines the subjunctive an *adjunctive* because it cannot convey any meaning alone. Something else must be added in the discourse: "Subiunctivus sive adiunctivus ideo dictus, quod per se non exprimat sensum, nisi insuper alius addatur sermo, quo superior patefiat" [The subjunctive, also called adjunctive, because alone it does not express any meaning if additional discourse that better discloses the meaning is not added] (GL I. p. 340. 24). The same stance is found in Cledonius: "Coniunctivus ideo dictus, quia solus sensum implere non potest" [The subjunctive is thus named because alone it can't express any meaning]. (K.V 16, 14)

The Latin grammarian Priscian further defines the syntactic environment which prompts the selection of the subjunctive mood and gives it some meaning. In particular, he claims that in addition to the conjunction or an adverb, the presence of another verb, the matrix verb of the main clause, is essential for the meaningful occurrence of subjunctive morphology: "Quartus [modus] est subiunctivus [...] qui eget non modo adverbio vel coniunctione, verum etiam altero verbo, ut perfectum significet sensum [...]" [The subjunctive is the fourth mood [...] which needs not only an adverb or a conjunction, but in truth, it needs also another verb, in order to have a complete meaning] (GL p. 424. 12).

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

A syntactic approach is essentially the one adopted by the Roman and Greek writers who see the indexing of the syntactic relation of subordination as the prime function of the subjunctive. Beyond that, the mood must rely on other lexical elements present in the close syntactic environment to fully acquire a meaning in context.

The rejection of the idea that the subjunctive acts as a full semantic marker is in opposition to the grammatical tradition of the Romance languages, such as Italian, French, Spanish and Portuguese.

In order to evaluate the viewpoint of the ancient grammarians in relation to the semantic accounts of the mood distribution in the Romance languages, the role of the subjunctive in Latin must be analyzed and compared to its function in the descended idioms. This is naturally possible considering Classical Latin, i.e. the standard and literary variety of Latin of which we have the most extensive written documentation through the works of the classical authors of the Roman Republic and of the Roman Empire. Its grammar is the most comprehensively reported and most taught today of the varieties of Latin (Teuffell, 1873; Bennett, 1946; Pinkster, 1990; Janson, 2004; Allen & Greenough, 2006; Sihler, 2008; Baldi & Cuzzolin, 2010). Additionally, we can rely on the sources on the *Sermo Vulgaris*, which encompasses the colloquial, non-standard varieties of Latin. The latter are to be considered the most direct ancestor(s) to the Romance languages (Diez, 1882; Wright, 1982; Grandgent, 1907, 1991; Williams, 1962; Lloyd, 1987; Malkiel, 1992; Price, 1998; Bonfante, 1999; Herman, 2000; Banniard, 2008; Ledgeway, 2012; Ledgeway &

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Maiden 2016; Maiden, 2014). Pinkster (2015), in particular, observes that some sources give us a certain access to spoken Latin: “Plautus’ comedies, for example, are a (stylized) form of speaking and so are the conversational parts of Cicero’s dialogues. Petronius’ *Cena Trimalchionis* is another example of stylized conversation. We also possess recorded texts (or texts based on protocols), like the *Passio sanctorum Scillitanorum* (the trial took place in AD 180) and the *Acta of the conference at Carthage in 411*” (p. 3). Such a collection of texts naturally falls short of the standards required nowadays for linguistic corpora that may be deemed a representative and balanced sample of a certain language variety (Sinclair, 2005). There is a submerged spoken Latin which is not accessible to us through the sources that we have and whose features can only be indirectly recovered through the contrastive analysis with the Romance languages. In addition, as already stated, a comparative analysis of different stages of language development throughout the centuries helps us piece together the larger picture of the evolution of morphological systems. These must be considered in relation to each other in the task of mapping the relevant space of meaning.

Latin had three moods: the indicative (*modus indicativus*), the subjunctive (*modus subiunctivus* or *coniunctivus*), and the imperative (*modus imperativus*) (Pinkster, 2015, p. 56). The Latin subjunctive derives from the Proto-Indo-European optative mood which was used for “more saliently conditional, doubtful or wished-for future events” (Sihler, 2008, p. 592). The Proto-Indo-European subjunctive, like the

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

optative, was an irrealis mood but with more “feeble modal force”. It used to indicate “a future event anticipated with some slight reservation on the part of the speaker —the equivalent of ‘I suppose’ or ‘in that case’”. In Vedic, the subjunctive [was] most often a simple future, occasionally something little different from an imperative” (Sihler, 2008, p. 592). This attenuated modal force, according to the author, made it possible for the Proto-Indo-European subjunctive to evolve into the Latin simple future tense.

4.2 The subjunctive mood in unembedded utterances

The accounts of the subjunctive in Latin testify to its use in unembedded utterances, some of which are no longer possible or are far less productive and restricted to a few, more constrained syntactic structures in the Romance languages.

Generally, Pinkster (2015) presents the subjunctive as a grammatical system available in Latin to mitigate the illocutionary force of an assertion or a question compared to the indicative: “the subjunctive may be used as a milder way of putting things” (p. 310) either for reasons of politeness or uncertainty. In (131) is a polite assertion and in (132) a polite interrogative where the question appears less direct. The sentences are taken from Pinkster (2015, p. 310 and p. 341).

(131) Hoc vero sine ulla dubitatione confirmaverim . . .

This, truthfully without any hesitation confirm.1.SG.PRF.SUBJ

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

“This, however, I would affirm without any hesitation . . .” (Cic. Brut. 25 in Pinkster, 2015, p. 310)

(132) Quid est? / # Possis, si forte accubantem tuum virum conspexeris
/ cum corona amplexum amicam, si videas, cognoscere?

What be.3.SG.PRS.IND? / # can.2.SG.PRS.SUBJ, if maybe
lie.PART.PRS your husband see.2.SG.PRF.SUBJ / with garland
hug.PP, if see.2.SG.PRS.SUBJ, know.INF

“What’s the matter? # If you were to see your husband lying there
with a garland
embracing his mistress, could you recognize him if you were to
see him?” (Pl. As. 877–9)

In imperative sentences, it is more difficult to make a general statement about how the subjunctive modulates the directive illocutionary force. Pinkster (2015, p. 351) notes that, compared to the imperative mood, it is used more frequently in binding directives, those that formally oblige the addressee to fulfill the content of the command or the prohibition.

Beyond those very abstract generalizations that cannot really be employed in the interpretation of the subjunctive in each specific clause, Latinists speak of different *functional categories* of the subjunctive. Indeed, scholars like Pinkster (2015, p. 390) give up on the notion that a *common denominator* for the functions of the subjunctive in the main clause can be traced.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

The most relevant functional categories of the subjunctive treated in grammars of Latin are listed and discussed in the following where they are compared to the grammar of the Romance languages.

4.2.1 Volitive subjunctive

The *volitive* or *deontic subjunctive* is considered an umbrella term in grammars of Latin and Greek. It encompasses a series of functions that pertain to the sphere of the subject's will. Basile (2001) among others speculates that it is the original value of the subjunctive from which the *epistemic* or *potential subjunctive* later developed (recall that Kempchinsky [2009, p. 1796] expressed a similar view).

Within the volitive domain, grammars usually distinguish the following categories:

- a. Hortatory subjunctive also called exhortative subjunctive (*coniunctivus adhortativus*)
- b. Jussive subjunctive (*coniunctivus iussivus*)
- c. Prohibitive subjunctive (*coniunctivus prohibitivus*)

Sometimes, the categories are collapsed together, e.g. the *hortatory-jussive subjunctive*, or sorted in different sets: e.g. *the concessive* and the *deliberative subjunctive* are presented under the volitive category by some Latinists. In other grammars, the deliberative subjunctive is often included within the potential subjunctive (Pinkster, 2015, p. 489).

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

- a. The *hortatory subjunctive* encodes an exhortation. With this function, Bennett (1946, p. 176) notes that the subjunctive is used most often with the first person plural of the present and has the value of “a proposal for some joint action of the speaker and the addressee” (Pinkster, 2015, p. 497). Nonetheless, first-person singular forms with hortative meaning were also used. Examples are sentence (102) reported here again in (133) and sentence (134).

(133) Amemus patriam.
love.1.PL.PRS.SUBJ country.ACC
“Let us love our country.” (Cicero, *Orationes*, *Pro Sestio*, in
Cicero [1895, p. 70] reported in Tantucci [1962, p. 244])

(134) Ne difficilia optemus neve inania consecemur
Not difficulties desire.1.PL.PRS.SUBJ and not vain things
pursue.1.PL.PRS.SUBJ
“Let us not desire difficult things and let us not pursue vain
things” (Cicero reported in Tantucci, 1962, p. 244)

- b. The *jussive subjunctive* expresses a command. According to Bennett (1946, p. 177) with this function it is most frequently used in the third person singular or plural. Examples are sentence (104) reported again here in (135) and sentence (136).

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(135) Arma deponat

Weapon.PL.ACC lay.3.SG.PRS.SUBJ

“He must lay down the weapons!” (with imperative value: *Lay down the weapons.*) (Cicero, Philippics 3-9, Volume 1, in Manuwald, 2012, p. 200)

(136) Dīcat

Tell.3.SG.PRS.SUBJ

“Let him tell” (in Bennett, 1946, p. 177)

It is much less frequent in the second persons due to the competition with the present and future imperative mood (Pinkster, 2015, p. 498). According to Tantucci (1962, p. 244), in these rare uses, rather than a command, it conveys the illocutionary force of a request or a plea and it is found in more familiar registers. Pinkster on the other hand reports that the subjunctive is selected for binding commands.

- c. The *prohibitive subjunctive* is used together with negation particle *ne* to express a prohibition with the second and third person singular and plural. Like the jussive subjunctive in affirmed utterances, this type of subjunctive is more common in the third persons than in the second.

(137) Ne attingant rem publicam.

NEG touch.3.PL.PRS.SUBJ res.SG.ACC publica.SG.ACC

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“They must not touch the *res publica*.” (with imperative value:

Do not touch the res publica.)

(*Orationes*, Pro Sestio, in Cicero [1856b, p. 569] reported in Tantucci [1962, p. 244])

A locution that uses the imperative of the verb *cavere*, *cave* (lit. be cautious!) with the subjunctive is more common in the second person. Bennett reports among others an example adapted here in (138).

(138) *Cave ignoscas*.

Avoid.2.SG.PRS.IMP forgive.2.SG.PRS.SUBJ

“Do not forgive!” (Cicero, Pro Ligario 14 also in Bennet, 1946, p. 177)

All the above illocutionary values of the subjunctive, i.e. exhortations, commands, requests, and together with negation particles, prohibitions, are considered expressions of a unique semantic domain by some grammars, one pertaining to the directive or deontic modality. In the Romance languages, this semantic domain is mainly conveyed by the imperative mood. In Latin, however, the imperative mood system had only two persons, the second singular and plural. All the other persons were supplemented by the subjunctive. Nowadays the subjunctive has largely been substituted by the imperative mood in the main locution.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Italian retains a relic of its original functions: the polite forms of address grammaticalized in the third person singular *lei* “You” and plural *loro* “You” require a subjunctive form for expressing polite exhortations, orders or prohibitions in a more formal setting. In this use, they supplement the morphological gaps in the system of the imperative which lacks third-person forms (see utterances in (139 a and b)). Table 4.1 presents the imperative paradigm with the morphological gaps supplemented by the subjunctive.

(139)

- a. La prego, si sieda. vs. Siediti per piacere.

PRO ask.1.SG.PRS.IND, REFL.PRO sit.3.SG.PRS.SUBJ vs.
sit.2.SG.PRS.IMP to please.

“Please, take a seat.” [formal address] vs. “Sit down, please
[informal you].”

- b. Se ne vadano! vs. Andatevene!

REFL.PRO ADV go.3.PL.PRS.SUBJ vs. go.3.PL.PRS.SUBJ-REFL.PRO-
ADV

“Go away!” [formal address to a plurality] vs. “Go away!”
[informal, plural *you*]

As such, Italian exhibits asymmetries in the morphological resources used to convey the very same function, with two different moods concurring to express different grammatical persons.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Table 4.1

Imperative paradigm of Italian verb parlare “talk”

Imperative		
1 SG	-	
2 SG	<i>parla</i>	
3 SG		<i>parli</i> (SUBJ)
1 PL	<i>parliamo</i>	
2 PL	<i>parlate</i>	
3 PL		<i>parlino</i> (SUBJ)

In b., I discussed how in Latin the subjunctive was also used to give a command to a third-person referent, i.e. a referent that is not addressed directly (he, she, they) (see again (135), (136)). French and Italian lack third-person forms of the imperative. They both present a construction based on the fixed combination of the complementizer *que, che* “that” and a third-person subjunctive, two elements that have grammaticalized together. This construction is used for orders or prohibitions directed to one or more referents who are generally not present (140 a and b). There are a number of more or less formulaic phrases which make use of this syntactic construction. An example is given in (140c). where the order addresses a plurality.

(140)

- a. Fr. *Qu’ils s’en aillent!* It. *Che se ne vadano!*
 Fr. That they REFL.PRO ADV go.3.PL.PRS.SUBJ It. That they
 REFL.PRO ADV go.3.PL.PRS.SUBJ.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“They must go away!”

- b. Fr. *Que personne ne se sente blessé de ne pas l’être!* It. *Che si faccia gli affari suoi!*

Fr. That nobody NEG REFL.PRO feel.3.SG.PRS.SUBJ hurt to NEG be it! It. That REFL.PRO do the affairs his!

“Do not interfere!”

- c. Fr. *Que personne ne bouge!* It. *Che nessuno si muova!*

Fr. That nobody NEG move.3.SG.PRS.SUBJ! It. That nobody REFL.PRO move.3.SG.PRS.SUBJ!

“Nobody move!”

Additionally, in Italian, the imperfect subjunctive can substitute the present subjunctive in main clauses with jussive value either introduced by *che* or not. This use mainly characterizes the Italian of the Central and Southern part of the Italian peninsula (Telve, 2011). See utterances (141) and (142).

- (141) *Che si facessero venire un’idea adeguata al periodo che stiamo vivendo!*

That REFL.PRO make.3.PL.PST.SUBJ come.INF an idea adequate to-DET time that be.1.PL.PRS.IND live.

“They must come up with an idea suitable to the time that we are living in!” (blog.repubblica.it, ItWac)

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(142) (Che) parlasse chiaramente!

(That) speak.3.SG.PST.SUBJ clearly!

“S/he must speak clearly!”

These uses of the subjunctive in the main locution are however rather infrequent in speech, if one excludes the subjunctive forms acting as imperatives in formal settings. In the Italian C-ORAL-ROM corpus, only 8.3% of the subjunctives that occur in independent clauses exhibit this jussive function. The percentage is considerably reduced if one considers all the subjunctive occurrences in the corpus, including the forms appearing in subordinate clauses (2.4%). Only one instance involves a subjunctive imperfect in a linguistic setting characterized by the use of a central Italian dialect of the Umbria region. Examples of utterances extracted from the C-ORAL-ROM are reported in (143) and (144).

(143) Non vi sembri una sottigliezza così, semplicemente di blasone.

Not you.PL.DAT seem.3.SG.PRS.SUBJ a subtlety that, simply of blazon.

“It must not seem to you a trivial matter of rank.” (inatte01, C-ORAL-ROM)

(144) Ci sia ordine nel capire che tipo di cliente hai

ADV be.3.SG.PRS.SUBJ order in-DET understand.INF what type of client have.2.SG.PRS.IND

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“There must be order in understanding what type of client you have.” (ilfamcv16, C-ORAL-ROM)

4.2.2 Deliberative subjunctive

The *deliberative subjunctive* (*coniunctivus deliberativus*) also called *dubitative subjunctive* was used in direct interrogative utterances or exclamations to express a doubt, an uncertainty, or indignation. The present was used for queries concerning the present or the future (see (145)) while the imperfect tense encoded a doubt concerning the past (see (146)). The negation particle for this type of subjunctive is *non*. (Tantucci, 1962, pp. 244-245; Pinkster, 2015, p. 490).

(145) Quid agam, iudices? Quo accusationis meae rationem conferam?

Quo me vertam?

What do.1.SG.PRS.SUBJ, judges? Where accusation.GEN my.GEN
reason.ACC bring.1.SG.PRS.SUBJ? Where I.ACC
direct.1.SG.PRS.SUBJ?

“What should/could I do, judges? Where should/could I bring the motive of my accusation? (Where should I begin to build my accusation?) Where should/could I direct myself?” (Cicero, Ad Verrem, II, 5, 1-4)

(146) Haec cum viderem, quid agerem, iudices?

Those CONJ see.1.SG.PST.SUBJ, what do.1.SG.PST.SUBJ, judges?

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

“Having seen all that, what should I have done, judges?” (Cicero, Pro Sestio, 42 in Pinkster, 2015, p. 490)

Italian and other Romance languages have shifted to other morphological and lexical strategies to carry out this particular function of the Latin subjunctive. Questions expressing doubt are usually formulated with modal verbs and forms of the indicative mood like the present or the future or the conditional mood. A few examples are given in (147) and in (148). The conditional and the indicative future paradigms constitute a morphological evolution of Common Romance with respect to Latin (Wolper, 1874; Adams, 2013; Zafiu, 2013). As a matter of fact, the modern synthetic future of the Romance languages arose from the morphological reduction and fusion of infinitive forms with Latin present tense of *habere* “to have” (Valesio, 1968; Clancy, 1975). This periphrasis is assumed to have originally expressed the meaning of obligation or necessity in terms of predestination (i.e. “what is to happen”) and in time simply came to encode futurity with an intermediate period in which both meanings coexisted (Grandgent, 1907, p. 57; Valesio, 1968; Benveniste, 1968; Clancy, 1975, p. 547). An analogous development led to the formation of the conditional, a paradigm that is completely absent in Latin. It arose from the reduction and fusion of infinitive forms with past forms of *habere* and was eventually used to encode anterior futurity (i.e. the future in the past) or conditionality. Those meanings are still conveyed by the conditional today (see Coleman [1971] reviewed in Graham [2015, pp. 12-13]). In

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

Latin they were expressed by the pluperfect subjunctive according to Grandgent (1907, p. 58). For a more detailed history of the formation of the conditional and the future paradigms from the Vulgar Latin to Common Romance see Clancy (1975).

(147) It. Cosa dovrei dire? Cosa avrei potuto dire? Cosa dirò?

It. “What should I say?” “What could I have said?” “What will I say”

(148) Sp. ¿Que debo hacer? ¿Que debería decir?

Sp. “What should I do?” “What should I say?”

4.2.3 Concessive subjunctive

The *concessive subjunctive* (*coniunctivus concessivus*) is used in Latin in the main clause to indicate that something, a fact or an event, is “granted or conceded for the sake of the argument” (Bennet, 1946, p. 178). Both the present and the perfect subjunctive were used productively. The former for concessions regarding the time of speech (see (149)) and the latter for admissions concerning the time anterior to the time of speech (see (150)) (Pinkster, 2015, p. 509; Tantucci, 1962, p. 249). Pinkster notes that in such clauses there is a tendency to front the verb or to put it “as early as possible in the clause”. Also, he remarks that it is sometimes difficult to interpret such subjunctives and distinguish between a jussive and a concessive meaning.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(149) Haec [...] sint falsa sane; invidiosa certe non sunt

These [...] be.3.PL.PRS.SUBJ false certainly; odious certainly not are.

“Granted that these things are false, they are certainly not hideous.” (Cicero1856a, p. 109; also in Tantucci, 1962, p. 249)

(150) Fuerint cupidi, fuerint irati, fuerint pertinaces. Sceleris vero crimine furoris parricidi [...] liceat multis aliis carere.

Be.3.PL.PRF.SUBJ greedy.PL.NOM be.3.PL.PRF.SUBJ
wrathful.PL.NOM, be.3.PL.PRF.SUBJ obstinatePL.NOM.
Villainy.GEN true.ABL accusation.ABL revolt.GEN treason.GEN
[...] be-lawful.3.SG.PRS.SUBJ many. others lack.INF.

“Granted that they were greedy, granted that they were wrathful, granted that they were obstinate, but let the many others be lawfully exempted from the accusation of villainy, of revolt, of treason.” (Cicero, Pro Ligario, 18 in Pinkster, 2015, p. 510)

In the Romance languages, this concessive value is generally no longer available in main clauses but it is expressed in subordinate utterances instead. The latter are usually introduced by subordinating elements like conjunctions and other lexical elements that overtly express the meaning of a concession: for instance It. *sebbene*, *nonostante*, *benché* “although”, *anche se* “even if”, *comunque vada* “however it goes”. The use of the subjunctive in the subordinate clause with concessive meaning is variable and largely depends on the

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

subordinating conjunction or the construction used (see the mood distribution with the French conjunctions *bien que* “although” and *même si* “even if” in (151 a and b) or the Italian conjunctions *benché* “although” and *anche se* “even if” in (152 a and b).

(151)

- a. **Bien qu'il soit issu** d'un milieu relativement modeste, Fernando est initié dès le plus jeune âge aux joies du karting.
Although COMPL-he be.3.SG.PRS.SUBJ originated PREP-an environment relatively modest, Fernando is initiated since the most young age to games of carting.

“Even though he comes from a relatively modest environment, Fernando is initiated at a very young age to the karting game.”

(flactu.ns7-wistee.fr, FrWac)

- b. Elle sent fort l'oignon, **même si elle est** pas allée dans la cuisine.

She smells strongly DET-onion, even if she be.3.SG.PRS.IND NEG gone to the kitchen.

“She smells the onion, although she didn’t go to the kitchen.”

(152)

- a. È un progetto che ha dell’innovativo **benché** la costruzione di strumenti musicali **sia** qualcosa di molto antico. (imedts05, C-ORAL-ROM)

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Is a project that has PREP-innovative although the construction of musical instruments be.3.SG.PRS.SUBJ something PREP very ancient.

“It is an innovative project although the construction of musical instruments is a very ancient practice.”

b. Si sentiva forte e bella **anche se era** piccola piccola.
(blog.libero.it, ItWac)

(She) REFL.PRO felt strong and beautiful even though be.3.SG.PST.IND little little.

“She felt strong and beautiful even though she was very little.”

In addition, in informal speech, mood variation can be encountered with the same subordinating element or phrase (see the indicative/subjunctive distribution with Spanish conjunction *aunque* “although” in (153 a and b) and Italian conjunction *nonostante* “although” in (154 a and b)). Certain concessive formulations exhibit a marked word order with verb fronting very similar to Latin: examples are Italian *ammesso che* “although” (see utterance (155)), *posto che* “although”, *sia pure che* “although”. These locutions are quite fixed. Some of them retain subjunctive forms, such as: *sia pure che* “although”, *male che vada* “however badly things go”, *checcché si dica* “whatever people may say”. They are discussed along with others in Serianni and Castelvechi (1989, p. 601).

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(153)

- a. Las tormentas [...] han traído agua, aunque haya sido poca. (19-07-23 ES, Cadena SER, Corpus del Español)

The storms [...] have brought water, although
have.3.SG.PRS.SUBJ been little

“The storms [...] have brought water, although it has been
little.”

- b. Su temporada 2019 se ha visto perturbada por varias caídas, aunque ha logrado terminar quinto [...].

His season 2019 has seen disrupted by various falls, although
have.3.SG.PRS.IND managed end fifth [...]

“His season 2019 was disrupted by various falls, although he
managed to end the race in fifth place [...].” (19-0-7--31 ES,
Diario Palentino, Corpus del Español)

(154)

- a. Nonostante hanno ideologie diverse lavorano insieme.
(filmup.leonardo.it, ItWac)

Although have.3.PL.PRS.IND ideologies different
work.3.PL.PRS.IND together.

“Although they have different ideologies, they work together.”

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

b. Non possono avere un libretto sanitario nonostante abbiano un regolare rapporto di lavoro. (stranieriinitalia.it, ItWac)

They can't have a book health although have.3.PL.PRS.SUBJ a regular relationship of work.

"They can't have health insurance even though they are regularly employed."

(155) Ammesso che propone e pone gli stessi obiettivi, deve determinare una politica chiara.

Admitted that propose.3.SG.PRS.IND and set.3.SG.PRS.IND the same objectives, must.3.SG.PRS.IND determine a politics clear

"Although it proposes and sets the same goals, it should determine a clear policy." (inatps03, C-ORAL-ROM).

Moving away from the domain of volition, Latinists have identified additional semantic values of the subjunctive in the main clause. The categories usually treated are listed and briefly discussed in the following passages.

4.2.4 Optative subjunctive

The *optative subjunctive* (*coniunctivus optativus*) is used for expressions with the illocutionary force of wishes. In this semantic domain, Latin employs all the tenses:

The present and the perfect tense encode wishes that can be realized in the present or in the past respectively. See utterances (156) and (157).

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

The imperfect and the pluperfect tense express *irrealis* wishes, those that are considered impossible to fulfill, either in the present (imperfect) or in the past (pluperfect). See utterances (158) and (159).

Pinkster (2015) notes that this type of expression with an optative value is often accompanied by special particles or adverbs like *utinam* (loosely translated “if only”) in all periods of Latin. Also, the optative subjunctive is sometimes directly preceded by other subjunctive forms like the present form *velim* “greatly wish”, *nolim*, “want”, *malim* “prefer” for realizable wishes and the imperfect *vellem* “wish greatly”, *nollem* “want”, *mallem* “prefer” for counterfactual wishes (Tantucci, 1962, p. 248).

(156) *Salvos sis, Mnesiloche. Salvom te advenire gaudeo.*

Safe be.2.SG.PRS.SUBJ, Mnesilochus. safe you.ACC arrive.INF
rejoice.1.SG.PRS.IND

“May you be well, Mnesilochus. I’m glad you’ve arrived safely.”

(Plautus, 1847, p. 689, in Pinkster, 2015, p. 505)

(157) *Utinam hinc abierit malam crucem.*

ADV. ADV. go.3.SG.PRF.SUBJ away evil gallows.SG.ACC

“May he have gone to be hanged!” (Plautus, 1847, p. 1315, in

Pinkster, 2015, p. 507)

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

(158) Homo hic ebrius est, ut opinor. - Utinam ita essem.

Man this drunk be.3.SG.PRS.IND, ADV think.1.SG.PRS.IND. - ADV
thus be.1.SG.IMPF.SUBJ

“This man is drunk, I think. # I wish I were.” (Plautus, 1847, p.
71, in Pinkster, 2015, p. 508)

(159) Quod utinam minus vitae cupidi fuisset.

That ADV less life greedy be.1.PL.PPRF.SUBJ

“Ah, if only we had been less eager for life!” (Cicero, 1790 in
Tantucci, 1962, p. 248)

The Romance languages of today exhibit to some extent the use of the subjunctive with an optative value in the main clause. Perhaps this is best preserved in Spanish among the Romance languages where subjunctive forms, past or present, are usually accompanied by certain particles like: *que*, *ojalá (que)*, *así*, *ya*, *quien* etc to encode wishes (see sentences (160), (161), (162)).

(160) Sp. Ojalá tuviera buenas noticias para todos

Sp. INTERJ have.1.SG.IMPF.SUBJ good news for everybody

“I wish I had good news for everybody” (19-04-25 US, Corpus
del Español: NOW)

(161) Sp. Ya hubiera yo querido un maestro así.

Sp. INTERJ have.1.SG.IMPF.SUBJ I wish.PTCP a teacher ADV.

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

“I wish I had had a teacher like that.” (16-07-13 EC, Corpus del Español: NOW)

(162) Sp. Mamá te mando muchos saludos, te amo, que tengas un buen día.

Sp. Mom you send.1.SG.PRS.IND many greetings, you.ACC love, that have.2.SG.PRS.SUBJ a good day

“Mom I send you many greetings, I love you, may you have a nice day.” (19-05-12 CL, Corpus del Español: NOW)

These optative utterances often exhibit a marked word order with the verb in subject position.

In contemporary Italian and French, such uses of the subjunctive in the main clause are more restricted. In French, especially, they appear to be mostly formulaic (see (163)), confined to written communication, formal and literary registers. The optative utterances may be introduced by *que* (see (164)) or by the subjunctive form of the modal verb *pouvoir* “can” to which the subject is postposed (165). Unlike Spanish and Italian, the past forms of the French subjunctive, the imperfect and pluperfect, have altogether disappeared from this and other semantic domains in both written and oral communication (Poplack, 1990, p. 18).

(163) Vive la France !

Live.3.SG.PRS.SUBJ DET France

“Long live France!”

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

(164) Que cela nous serve de leçon pour faire mieux à l'avenir.

That us serve.3.SG.PRS.SUBJ PREP lesson to do.INF better to DET-future.

“May that serve us as a lesson to do better in the future.”

(Wolfgang Schüssel, *europarl.europa.eu*)

(165) Puissent notre foi et notre indignation susciter des engagements individuels et collectifs.

Can.3.PL.PRS.SUBJ our faith and our indignation ignite.INF PREP-DET activism.PL individual and collective.

“May our faith and our indignation spur among us individual and collective activism.” (*cccb.ca*)

In Italian, optative utterances can in principle be formed with both the present and the past tenses of the subjunctive (see (165) which shows the common use of the modal verb *potere* “can” in this type of sentence and (167)). In the C-ORAL-ROM corpus, however, they constitute a tiny percentage of the subjunctives found in the main clause, 2.5%, and only 0.6% of all the subjunctives. Although the small number of occurrences does not allow any mapping of the distribution of the tenses in this domain, the past forms such as the imperfect and the pluperfect seem to be more frequently used than the present tenses, contrary to French. All the instances in the corpus exhibit the imperfect (168).

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

In Italian, as well, optative sentences tend to have a marked word order with the verb raised to subject position similarly to the ancient Latin structure from which they originated.

(166) Possiate trovare presto una soluzione!

can.2.PL.PRS.SUBJ find.INF soon a solution!

“May you soon find a solution!”

(167) L’avessi ascoltata!

PRO-have.1.SG.IMPF.SUBJ hear.PP!

“Had I only listened to her!”

(168) Ma tu, avessi una regola, Leonardo.

But you, have.2.PL.IMPF.SUBJ a rule, Leonardo.

“If only you had a rule, Leonardo.” (ifamd108, C-ORAL-ROM)

In addition, the analysis of the corpus data reveals the common use of a type of construction which employs the imperfect and pluperfect subjunctive and is directly derived from the optative utterance. The semantics of this construction has somewhat shifted from its original meaning although the source remains quite transparent. It is used as an intensifier: It emphasizes and gives emotional force to an assessment, an experience shared by the speaker. For example, *sapessi quanto è strana* (lit. “if only you knew how strange she is”) “she is really

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

strange”. In (169) and (170) I report a few of such occurrences in the C-ORAL-ROM.

(169) *Avessi visto che pensione bella*

Have.2.SG.IMPF.SUBJ see.PST.PTCP what pension beautiful

“The pension was so beautiful.” (ilfammn01, C-ORAL-ROM)

(170) *Sentissi come cantava*

Hear.2.SG.IMPF.SUBJ how sing.3.SG.IMPF.IND

“She sang so beautifully.” (ilfammn01, C-ORAL-ROM)

The unembedded optative utterance with a subjunctive form usually fronted, a vestige of the ancient Latin structure, is less productive in most Romance languages than what it must have originally been in Latin based on the accounts reviewed above. French and Italian, for one, show very sparse instances in speech corpora and do not productively employ all the tenses that were available in Latin with all the different semantic interpretations, both temporal and modal. In the Romance languages of today, the optative subjunctive in the main clause competes with other moods for the expression of wishes, like the indicative or the conditional. These moods are frequently used with verbs that lexically encode the optative illocutionary force. In Italian, they are the modal *volere* “want” (utterances (171) and (172)), *sperare* “hope” (utterance (173)), *augurare* “wish” (utterance (174)) or *piacere* “like” (utterance (175)). They may be followed by infinitive or

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

complement clauses that encode the content of the wish. Such constructions are considerably more common in the C-ORAL-ROM corpus. In this type of optative sentence, the subjunctive is confined to the embedded complement clause, where it is oftentimes discarded in favor of other moods, like the indicative (see (172)). See similar instances in French with conditional or indicative forms of verbs *vouloir* “want” (176), *aimer* “like” (177), (178)), *espérer* “hope” ((179) and in Spanish with verbs *desear* “wish”, *gustar* “like”, *querer* “want”, *esperar* “hope” (sentences (180), (181)). They may govern infinitive or complement clauses. In the latter, the subjunctive also appears variable (sentence (179) licenses an embedded indicative while in sentence (178) a subjunctive occurs).

(171) Vorrei che il tribunale vi riflettesse un attimo.

Want.1.SG.PRS.COND that the court PRO reflect.3.SG.IMPF.SUBJ a bit.

“I would like the court to reflect on it a bit.” (inatla03, C-ORAL-ROM)

(172) Non voglio che lei conosca Greta.

NEG want.1.SG.PRS.IND that she meet.3.SG.PRS.IND Greta.

“I don’t want her to meet Greta.” (ilfamd120, C-ORAL-ROM)

(173) Speriamo che non sia niente di grave.

Hope.1.PL.PRS.IND that NEG be.3.SG.PRS.SUBJ nothing serious.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“Let’s hope it’s nothing serious.” (imedsp01, C-ORAL-ROM)

(174) Mi auguro che non sia una fregatura.

pro wish.1.sg.prs.ind that neg be.3.sg.prs.subj a scam.

“I hope it is not a scam.” (inatpd03, C-ORAL-ROM)

(175) A me piacerebbe che non ci fosse questo comportamento schizofrenico.

PREP PRO like that NEG ADV be.3.SG.IMPF.SUBJ this behavior
schizophrenic

“I’d like for this schizophrenic behavior to stop.” (inatptd02, C-ORAL-ROM)

(176) Je voudrais qu’on en parle un peu plus longuement.

I want that-PRO PRO speak.3.SG.PRS.SUBJ a bit more extensively.

“I would like that we talk about it a bit more extensively.”
(ffammn104, C-ORAL-ROM)

(177) J’aimerais quand même pouvoir m’expatrier en Afrique

I-love.1.SG.PRS.COND ADV can-INF PRO-expatriate to Africa.

“I hope I can move to Africa.” (ffammn24, C-ORAL-ROM)

(178) J’aimerais que tu me fasses un gâteau

I-love.1.SG.PRS.COND that you me make.2.SG.PRS.SUBJ a cake.

“I would like you to bake me a cake.” (fnatpr03, C-ORAL-ROM)

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(179) J'espère que ça va être bien, quoi

I-hope-1.SG.PRS.IND that PRO go.3.SG.PRS.IND be.INF good, PRO
“I hope that it will be good.” (ffamvc05 , C-ORAL-ROM)

(180) No quiero que os echéis a llorar

NEG want.1.SG.PRS.IND that PRO throw.2.PL.PRS.SUBJ PREP
cry.INF
“I don’t want that you cry.” (efamcv03, C-ORAL-ROM)

(181) Me gustaría que este libro fuese escuchado en el más amplio
ámbito posible

PRO like.1.SG.PRS.COND that this book be.3.SG.IMPF.SUBJ hear.PP
in the most ample domain possible
“I would like for this book to be heeded in the most ample
domain possible.” (19-07-30 ES, Corpus del Español: NOW)

4.2.5 Potential subjunctive

The *potential subjunctive* (coniunctivus potentialis) in main declarative or interrogative clauses in Latin used to encode the possibility that an event or state of affair may occur according to the subject. The subjunctive present or perfect tense were employed with no substantial difference for a present or future possibility (see (183) and (184)) whereas the imperfect subjunctive was used to express possibility in the past.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“[The present subjunctive] is especially frequent in declarative clauses with first person forms of modal verbs, especially those of ‘wishing’, and verbs indicating a subjective statement” (Pinkster, 2015, p. 483) (see (182)). According to Bennett (1946, p. 179), such forms are the equivalent of the “may”, “should”, “would”, “can” and “could” potential in his terms. The “can” and “could” potential appear most commonly in the second person present or imperfect (184) (Tantucci, 1962, p. 245). Finally, Pinkster remarks that “third person forms often have an indefinite pronoun as their subject (*quispiam, aliquis* “someone”) or a generic third person plural (“they”) [...]” (p. 483) (see (185)).

(182) Quod tibi suadeam, suadeam meo patri.

PRO you.DAT advice.1.SG.PRS.SUBJ, advice.1.SG.PRS.SUBJ my dad.

“What I’m advising you I’d advise my own father.” (Plautus, 1847, p. 321 in Pinkster, 2015, p. 483)

(183) Mori nemo sapiens miserum duxerit.

Dying no wise wretched think.3.SG.PRF.SUBJ

“No wise person could consider death a wretched thing.” (Cicero, 1977, p. 102; in Tantucci, 1962, p. 245)

(184) [...] numquam postilla possis prendere.

Never afterwards can.2.SG.PRS.SUBJ take.INF

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

“[...] you could never get hold of him again.” (Plautus, 1847, p. 313 in Pinkster, 2015, p. 483)

In the Romance languages, the weakening of the modal force to express subjectivity and uncertainty in unembedded utterances relies on other mood forms like the conditional mood, the future present or perfect indicative or the present indicative, often in combination with the modal verb “can”. As an example, the utterance “someone might say”, which selects either a present or perfect subjunctive in Latin (see (185)), displays the conditional or the future or the present indicative with modal verb “can” in Italian (186), French (187) or Spanish (188).

(185) Lt. *dicat aliquis / dixerit aliquis*

say.3.SG.PRS.SUBJ someone / say.3.SG.PRF.SUBJ someone

“someone may say” (in Bennett, 1946, p. 179)

(186) It. Qualcuno direbbe (che) / Qualcuno dirà (che) / Qualcuno potrebbe dire (che) / Qualcuno potrà dire (che)

Someone say.3.SG.PRS.COND (that) / Someone say.3.SG.FUT (that) / Someone can.3.SG.PRS.COND say.INF (that) / Someone can.3.SG.FUT say.INF (that)

“someone may say (that)”

(187) Fr. On dirait (que) / D'aucuns pourraient dire / D'aucuns diront / On peut dire / On va peut-être dire (que)

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Someone say.3.SG.PRS.COND (that) / PRO.PL
can.3.PL.PRS.COND say.INF / PRO.PL say.3.PL.FUT / PRO
can.3.SG.PRS.IND say.INF / PRO go.3.SG.PRS.IND maybe
say.INF (that)
“someone may say”

(188) Sp. Alguien diría (que) / Se podría decir (que) / Algunos dirán
(que)
Someone say.3.SG.PRS.COND (that) / PRO can.3.SG.PRS.COND
say.INF (that) / Some say.3.PL.FUT (that)
“Someone may say”

As mentioned, in Latin the potential subjunctive figures prominently in direct interrogative sentences as well. Tantucci (1962) cites a few examples from Cicero’s writings reported here in (189) and (190).

(189) Quis dubitet quin in virtute divitiae sint?
Who doubt.3.SG.PRS.SUBJ that in virtue riches be.3.PL.PRS.IND
“Who could doubt that in virtue lies prosperity?” (adaptation
from Cicero in Tantucci, 1962, p. 245)

(190) Hoc tantum bellum quis umquam arbitraretur ab uno imperatore
confici posse?
This big war who ever think.3.SG.IMP.F.SUBJ from one
commandant conduct.INF.PASS can.INF?

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“Who would have imagined that such a big war could have been conducted by a single commandant?” (adaptation from Cicero’s *Orationes in* Tantucci, 1962, p. 245)

In Italian, the subjunctive has endured in interrogative utterances that express a possibility, a hypothesis or a wish. This particular type of construction is introduced by the subordinating conjunction *che* ‘that’. With the imperfect subjunctive the conjunction is not obligatory. According to Patota (2010), the subordinating element implies the existence of a matrix clause of the type *credi* “do you believe”, *credete* “do you believe”, *pensi* “do you think”, *pensate* “do you think” that governs the subjunctive clause and is not realized. Examples of subjunctive interrogatives are reported in (191), (192) and (193). No instance of subjunctives in direct interrogatives was found in the Italian C-ORAL-ROM. The mood competes with other verb forms in direct questions, like the future indicative or the infinitive.

(191) *Che sia incinta?*

That be.3.SG.PRS.SUBJ pregnant?

“Might she be pregnant?”

(192) *(Che) fosse incinta?*

That be.3.SG.IMPF.SUBJ pregnant?

“Might she be pregnant?”

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(193) Che abbiano voluto farla fuori?

That have.3.PL.PRS.SUBJ want.PP do.INF-PRO out?

“May have they wanted to exclude her?”

4.2.5.1 The conditional sentence

In addition to the contexts discussed above, in Latin the potential subjunctive was used frequently in the conditional sentence encoding a possible or a counterfactual scenario, both in the protasis and in the apodosis. The protasis is the subordinate utterance, the *si* clause expressing a condition or hypothesis. The apodosis is the main clause expressing the consequence of that condition. The protasis can sometimes be omitted because the condition is implicitly understood. In that case, the apodosis stands alone.

The present subjunctive encodes a present or future possibility (see (194)). The perfect tense is generally used only in the protasis. Until early Latin, the perfect encodes anteriority in relation to the apodosis (Tantucci, 1962, p. 395). From early Latin onwards, it loses this temporal value and becomes semantically equivalent to the present (Pinkster, 2015, p. 491).

In the apodosis, the present subjunctive is common especially in Early Latin until Cicero; from his work onwards, there is variation between the traditional subjunctive and the indicative present or future (Pinkster, 2015, p. 484).

(194) Si quis vos interroget [...] quid aliud respondeatis [...]?

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

If someons you.ACC ask.3.SG.PRS.SUBJ [...] what other
reply.2.PL.PRS.SUBJ?

“If someone asked you, what would you reply?” (Livius
Patavinus, 1842 in Tantucci, 1962, p. 395)

The imperfect and pluperfect subjunctive were used frequently in conditional utterances both in the protasis and in the apodosis for an impossible condition and therefore a consequence (apodosis) that is contrary to fact. The imperfect encodes an unreal scenario in the present (sentence (195)) and the pluperfect indicates an unreal scenario pertaining to the past (utterance (196)).

(195) Aliquanto amplius valerem, si hic maneres.

Somewhat more fare.1.SG.IMPF.SUBJ, if here stay.2.SG.IMPF.SUBJ
“I’d fare somewhat better if you were to stay here.” (Plautus,
1847, p. 185 in Pinkster, 2015, p. 495)

(196) Pol si habuissem, satis cepissem miseriarum e liberis.

INTERJ. If have.1.SG.PPRF.SUBJ, enough take.1.SG.PPRF.SUBJ
afflictions from children.

‘If I’d had any, I’d have had my share of afflictions from my
children (Plautus, 1847, p. 992 in Pinkster, 2015, p. 495)

As was generally the case for utterances with directive, optative and concessive illocutionary force discussed above, from Latin to the

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Romance languages Spanish and Italian, the subjunctive has progressively disappeared from the apodosis. However, it is still preserved in the embedded if-clause. The apodosis on the other hand requires the present or past conditional following the temporal agreement with the protasis.

In the protasis, a reorganization of the functions associated with the subjunctive tenses has taken place: the present and perfect subjunctive are no longer used. Both a possible and a counterfactual condition are conveyed with the imperfect and the pluperfect subjunctive instead. The imperfect subjunctive in sentence (197) and the pluperfect in (198) encode a possible and contrary-to-fact premise in Italian respectively. The Spanish imperfect subjunctive conveys a possible scenario in sentence (199) and a counterfactual one in (200) like the pluperfect in (201).

(197) Avresti difficoltà a lavorare se usassi il tuo cognome?

“Would you encounter (PRS COND) difficulties if you used (IMPF SUBJ) your surname?” (imedts08, C-ORAL-ROM)

(198) Se fossi stato più attento magari me ne sarei accorto.

“If I had been (PPRF SUBJ) more attentive, maybe I would have noticed (PST COND) it.”

(199) Si quisieras tener tanta plata como Sebastián Piñera, tendrías que trabajar durante 15 años.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“If you wished (IMPF SUBJ) to have as much money as Sebastián Piñera, you would have to (PRS COND) work for 15 years.” (18-10-15 CL, Corpus del Español: NOW)

(200) Si estuviera de acuerdo con Comisiones Obreras, pues no hubiera hecho esta pregunta.

“If I agreed (IMPF SUBJ) with the trade union, I wouldn’t have asked (PPRF SUBJ) this question.” (emedts01, C-ORAL-ROM)

(201) [...] si hubiese sido omnívora Andrea habría gastado 2.54 toneladas de CO₂ a el año

“[...] if she had been (PPRF SUBJ) omnivorous, Andrea would have consumed (PRF COND) 2.54 tons of CO₂ each year.” (19-07-13 ES, Corpus del Español: NOW)

The subjunctive is not the only form available in the protasis. The present indicative is supposed to encode conditions that are more likely to occur.

In French, unlike Spanish and Italian, the subjunctive has altogether disappeared from the protasis as well. A possible or counterfactual condition is expressed with the imperfect and pluperfect indicative instead. In the apodosis the standard norm requires the use of the conditional similar to Spanish and Italian (see sentences (202), (203)).

(202) Si on demandait aux Français [...], ils répondraient oui.

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

“If one asked (IMPF IND) the French, they would reply (PRS COND) yes.” (lunion.presse.fr, frWac)

(203) Si on avait pas été aux urgences, l’hémorragie aurait continué à se développer.

“If we had not been (PPRF IND) at A&E, the hemorrhage would have continued (PST COND).” (ffammn05, C-ORAL-ROM)

Leblanc (2009) reviews the research on the historical development of the tenses and moods in conditional sentences in French over the course of the centuries. She reconstructs the diachronic restructuring of the hypothetical sentence from Latin as follows:

In Old French, the various subjunctive tenses were still used in both clauses of conditional complexes (Brunot 1966; Wartburg 19469) at some point in Old French or Early Middle French, the use of the imperfect in the protatis and the conditional in the apodosis was introduced in hypothetical complexes relating to the present or the future (Brunot and Bruneau 1969). Throughout the period leading to modern French, the new system with asymmetrical forms is found, along with the system inherited from Latin, which uses the subjunctive in both clauses (Wagner 1939). Additional competition comes from mixed forms (Ellis 1986; Grevisse, 1998; Wagner and Pinchon 1965): the subjunctive in one clause and the imperfect or conditional in the other [...] (p. 1004).

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

The competition between a variety of mood and tense forms in conditional sentences in the period leading up to modern French can be taken as an accurate characterization of the Romance languages at present. Although one particular morphological schemata has come to represent the standard norm in each language, synchronic variation between different forms prevails in speech. For instance, in the French conditional utterances in (204) and (205), we see the use of the past conditional in both main and the embedded clause of the conditional complex.

(204) Si on aurait été moins con il ya 90 ans peut être on n'aurait pas connu le nazisme.

“If we had been (PST COND) less stupid 90 years ago, maybe we wouldn’t have known (PST COND) Nazism.”
(exupery.cowblog.fr, frWac)

(205) Si on aurait mis un autre acteur, le film ne serait pas aussi brillant.

“If another actor had been cast, the film wouldn’t have been that brilliant.” (moscovici.typepad.fr, frWac)

This is by no means restricted to French. In Italian, the same mood and tense forms in both clauses of the conditional period are frequently encountered in speech. Such forms are considered non-standard or colloquial and are often sanctioned within the linguistic community due to the cultivation of a prescriptive norm (as was seen in chapter 2). In

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(206), (207) and (208), I present some examples with the imperfect indicative, a form that is especially common in speech to indicate past contrary-to-fact scenarios; in (209), (210) the conditional present appears in both clauses, a slightly more marked variant encoding a present possibility.

(206) Se io chiudevo con due carte te sapevi che avevo il tris in mano
“If I had closed (IMPF IND) with two cards, you would have known (IMPF IND) that I had three of a kind in my hand.”
(ifamcv14, C-ORAL-ROM)

(207) Se non ero sposata non potevo la potevo fare.
“If I hadn’t been married (PPRF IND) I couldn’t have done (IMPF IND) it.” (ilfamcv18, C-ORAL-ROM)

(208) Se ero da sola poi alla fine decidevo di fare altro.
“If I had been alone (IMPF IND), I would have decided (IMPF IND) to do something different.” (ilfamcv21, C-ORAL-ROM)

(209) Se avreste un briciolo di dignità scioperereste ad oltranza.
“If you had (PRS COND) a bit of dignity, you would strike (PRS COND) indefinitely” (fiorentina.it, itWac)

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

(210) La tua indignazione potrebbe aumentare, se verresti a sapere quanto guadagna un usciere impiegato del Ministero degli Affari esteri.

“Your indignation would increase (PRS COND) if you found out (PRS COND) how much an employee of the Ministry for Foreign Affairs earns.” (lafabbricadelprogramma.it, itWac)

The practice of repeating the morphological form already used in one clause in a syntactically related one has been referred to as *mood/tense harmony* in the literature. It has been discussed by a number of scholars in relation to conditional sentences and more generally as a phenomenon of the main and embedded clause in the Romance languages. It is considered a general trend in oral communication, one that prompts morphological change. LeBlanc (2009) reviews it as follows:

According to Damourette and Pichon (1936), the use of the conditional in both members of the hypothetical complex reflects a tendency for popular French to level out morphological asymmetries such as that of the prescriptive rule. Harris (1986) mentions a movement towards parallelism in many Romance languages. This tendency does not apply only to the protasis-apodosis relation in Romance languages or in colloquial French. Others have reported similar effects for other linguistic variables or dialects (Sankoff & Laberge 1978; Poplack 1980; Weiner & Labov 1983; Pereira-Scherre & Naro

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

1991). This tendency may be a good candidate for a universal rule of spoken language [...]" (p. 1014).

However, the above-mentioned morphological patterns that respond to the principle of mood and tense harmony are not the only ones competing with the prescribed forms. Also, mixed morphology between the matrix and the if-clause is found in various corpora. See the conditional clauses (211) and (212) exhibiting the imperfect indicative in the protasis and the past conditional in the apodosis in Italian.

(211) Se lo sapevo prima, non avrei speso altri soldi per il chip.

"If I had known (it) before, I wouldn't have spent other money on the chip." (cdcopy.it, itWac)

(212) Io quel film, se lo sapevo, non l'avrei nemmeno scaricato.

"If I had known (it) (IMPF IND), I wouldn't even have downloaded (PST COND) that film." (hackerjournal.it, itWac)

4.2.6 Conclusion

Reviewing the occurrence of the subjunctive mood in the main clause in Latin reveals how, in comparison with the ancestor, the subjunctive generally has more restricted domains of application in the unembedded locution in the Romance languages. In the contexts where it is still used nowadays, often not all the tenses of the paradigm can be employed as they were used in Latin (see again the protasis of conditional

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

sentences), or they have come to encode different functions: for instance, in certain constructions, the paradigm is so organized that the present and imperfect tense no longer index a specific point in time, whether present or past, in relation to the time of speech. Instead, they change the modal force of the utterance. Also, whereas in Latin the subjunctive could stand alone in the main clause, in Italian, French or Spanish it is often accompanied by the conjunction “that” which usually introduces subordinate clauses (see again the subjunctive with optative or directive illocutionary force, or the subjunctive in interrogative clauses in Italian).

Finally, considering the frequency of use and the productivity reported in the corpora of the Romance languages, the subjunctive in the main clause is for the most part infrequent and suffers from the competition of other morphosyntactic strategies. Some of its uses in the main locution are found in formulaic phrases. In particular, a quantitative analysis of the Italian C-ORAL-ROM corpus reveals that, of the overall 1,148 occurrences of the subjunctive, only very few are to be found in the main clause. Among those, there are mainly the subjunctives that act as imperative forms addressed to a direct interlocutor in a formal setting. If those instances are excluded, only 5.8% of all the subjunctives occur in unembedded utterances.

4.3 The subjunctive mood in subordination

In light of the fact that the subjunctive had a wider range of functions in the main clause in Latin in comparison to the Romance languages, it

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

might seem puzzling that Roman grammarians used to present it as the mood of the hypotaxis, largely dependent on the conjunction, and that they explicitly rejected the notion that it had an independent meaning. This rejection is clearly in opposition to the grammatical tradition of the Romance languages, such as Italian, French, Spanish and Portuguese shaped by the numerous efforts of grammarians and linguists to either find a core semantic value of the subjunctive or discover and analyze the most important ones.

In order to understand the position of the Roman grammarians, one must analyze the relative distribution of the subjunctive in the main and subordinate clause and trace its development from early to late Latin.

Pinkster (2015, p. 390) reports that out of the 368,463 verb forms found in the corpus of Classical Latin texts, LASLA (Denooz, 1978, 2004; Dee, 2002; Longrée, 2009), the most frequent forms are indicatives (48.2%) followed by subjunctives (16.2%), infinitives (15.6%), participles (14.5%), imperatives (2.6%), gerundives (2%), and gerunds (0.8%). The subjunctive is thus the second most frequent mood, even though its occurrences are far fewer than the indicative.

The scholar also looks at the mood distribution in a comedy by Plautus *Amphitryon*, on the grounds that the comedy genre best reflects oral interaction and can thus give a more accurate picture of how mood systems were employed in oral communication. He observes that in this particular sample overall the indicative is the most frequent mood. Nonetheless, the imperfect subjunctive occurs more frequently than the

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

imperfect indicative and the pluperfect subjunctive is as frequent as the pluperfect indicative (Pinkster, 2015, p. 391).

Moving beyond absolute frequencies, he examines the relative distribution of the two moods in main and embedded clauses in Plautus's comedy *Amphitryon* and in Caesar's *Bellum Gallicum*, two very different genres, the former is constituted by dialogues and monologues while the latter is a third-person narrative in prose form. He finds that most indicatives occur in unembedded clauses. On the contrary, the subjunctive is most present in embedded clauses: only 18% of all subjunctives in Plautus *Amphitryon* and less than 1% of the subjunctives in Caesar's *Bellum Gallicum* are found in unembedded clauses (p.393). He concludes that "The indicative is for the most part a main clause mood, the subjunctive a subordinate clause mood [...]" (p. 393).

From the language samples evaluated, the subjunctive emerges as the mood of subordination in Latin. In the embedded clause, its special relation with the subordinating conjunctions *ut* and *ne* is highlighted by more recent grammars of Latin. See among others Ledgeway (2012, pp. 12-13) who introduces the two conjunctions as subjunctive complementizers. Those accounts strengthen the position of the ancient Roman grammarians who presented the mood as essentially dependent on other elements of the clause, like the conjunction.

Pinkster (2015) reflects that already in Latin, beyond the main clause, "the use of the subjunctive in subordinate clauses can be better described as a gradient between the two poles '(fully) semantic' and

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

‘(fully) grammatical’” (p. 617) with cases at the extreme end of the grammatical pole where the mood must be understood as a mere sign of subordination:

The subjunctive mood is not the only phenomenon in Latin that shows semantic alongside purely grammatical use: cases and prepositions are other well-known instances [...]. For the hearer/reader, the context in which a subjunctive form occurs will normally provide sufficient information to determine whether the subjunctive form must be interpreted as semantic or as purely grammatical. It is likely, for example, that a Roman did not feel the need to discover the semantic contribution of a subjunctive form in an indirect question like the one in (a). In a context where a governing cognition verb (*scis* ‘you know’) is used in combination with an indirect question introduced by the question word (*ubi* ‘where?’), there simply was no element of choice involved. The same applies to (b), in which the ‘order’ predicate *iubet* is used in combination with the subordinator *ut*.

- a. *Eho, Parmeno mi, nostin’ et scis ubi siet?*

Hey, Parmeno my, know.2.SG.PRS.IND and
know.2.SG.PRS.IND where be. 3.SG.PRS.SUBJ

“Hey, my dear Parmeno, you know her? You know
where she is?” (Ter. Eu. 351, in Pinkster, 2015, p. 390)

- b. *Telobois iubet sententiam ut dicant suam.*

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

Teloboians.DAT order.3.SG.PRS.IND decision.ACC CNJ
tell.3.PL.PRS.SUBJ his.

“He orders them to tell the Teloboians his decision.” (Pl.
Am. 205) (p. 390)

4.4 The diachronic evolution of the subjunctive: from the main clause to the subordinate clause

A number of Latinists have speculated that the grammatical uses of the subjunctive in the subordinate clause originated from semantic ones in the main clause. More specifically, according to their hypothesis, the subjunctive was employed first in the main clause. Main subjunctive clauses could appear as co-ordinates to other main clauses in paratactic sentences⁷ (see clause A with prohibitive illocutionary force normally selecting a subjunctive in Latin conjoined with clause B though coordination, the comma, in (213)). Once they were transformed into hypotactic structures, they retained the subjunctive (compare (213) with the hypotactic structure presented in (214). Here clause A appears as a subordinate complement clause governed by clause B) (Bennett, 1907, p. 223; Palmer, 1986, p. 132; Magni, 2010, p. 244).

(213) The praetor must not return to the army (clause A), the consul
announced (clause B).

⁷ Paratactic sentences are simple sentences made of main clauses joined together by means of coordinating conjunctions, e.g. *and*, *but*. They contrast with hypotactic sentences, complex sentences made up of clauses joined together by means of subordination (Butler, 2003).

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

(214) The consul announced (clause A) that the praetor must not return to the army (clause B).

This, according to Pinkster, is especially credible in subordinate clauses that semantically resemble the main clauses where the subjunctive occurs. For instance, the subjunctive in an *ut* clause governed by verb *volo-volere* “want” is semantically closely related to the subjunctive in independent imperative sentences (Pinkster, 2015, p. 619). The same can be said for all the clauses that are governed by verbs encoding the directive modality, verbs that express orders, exhortations, prohibitions, warnings, advice etc. As hinted previously, they are generally introduced by the conjunction *ut* or (*ut*) *ne*. Magni (2010, p. 245) notices that the complement clauses selecting a subjunctive can also appear without an introductory conjunction, further evidence, according to her, that these type of hypotactic sentences originated from paratactic structures (see (215)).

(215) *hodie uxorem ducas ut dixi volo*

Today wife.acc lead.2.sg.prs.subj as say.1.sg.prf.ind
want.1.sg.prs.ind

“I want you to get married, as I have said” (Ter. Andr. 418 in Magni, 2010, p. 245)

A semantic resemblance to the optative subjunctive can be traced in clauses governed by verbs that express wishes, desires, hope, prayers

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

but also fear (see Handford [1947, p. 89] for a more detailed explanation of how the verb *fear* fits into the optative category). They often lack the subordinating conjunction as well (see (216)).

(216) *Eveniant volo tibi quae optas*

Happen.3.PL.PRS.SUBJ want.1.SG.PRS.IND you.DAT that
wish.2.SG.PRS.IND

“I wish that those things you ask for come to you” (Plaut. *Persa* 293 in Magni, 2010, p. 246)

Many additional types of subordinate clauses that select the subjunctive are found to bear a similarity to the main clause subjunctives. For instance, complement clauses governed by verbs expressing doubt and uncertainty, or complement clauses introduced by *si* with conditional value have been related to the potential subjunctive in the main clause (see (217) and (218) presented in Magni [2010, p. 246]). Moreover, the potential subjunctive is seen as the source for the use of the mood in reported speech or in consecutive clauses to encode potential results as well as actualized results.

(217) *si perrumpere possent conati*

If break through.INF can.3.PL.IMPV.SUBJ attempt.PP

“they (the Helvetii) tried to see if they could break through”
(Caes. *Gall.* 1,8,4 in Magni, 2010, p. 246)

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

(218) *miror [...] si quemquam amicum habere potuit*

wonder.1.SG.PRS.IND if some.ACC friend.ACC have.IND
can.1.SG.PRF.IND

“I wonder whether he (Tarquinius) could ever have had a friend”

(Cic. Lael. 54,2 in Magni, 2010, p. 246)

The deliberative subjunctive in direct questions expressing uncertainty, deliberation, indignation, surprise etc. is assumed to be the origin of the subjunctive in embedded indirect questions (see (219) originally in Pinkster [2015, p. 631] and Magni [2010, p. 247]). Magni (2010) notes that “in Early Latin the indicative was still retained in questions of fact” (p. 247) and that variation between the indicative and the subjunctive occurred in clauses governed by the same verb (an example is reported in (220)). Eventually, by 200 BCE, the subjunctive spread to all types of indirect questions (Handford, 1947, p. 173).

(219) *Loquere quid scribam, modo.*

Tell.2.SG.PRS.IMP what write.1.SG.PRS.SUBJ manner

“Just tell me what I should write” (Plaut. Bacch 745 in Pinkster, 2015, p. 631)

(220) *nunc cuius iussu venio et quam ob rem venerim dicam*

Now PRO.GEN command come.1.SG.PRS.IND and for what
purpose come.1.SG.PRF.SUBJ tell.1.SG.FUT.IND

4 *A diachronic analysis of the subjunctive mood from Latin to the Romance languages*

“now I will tell you by whose command I come and for what purpose I have come” (Plaut. Amph. 17 in Magni, 2010, p. 247)

Aside from the instances illustrated above and a few others, there are a variety of cases where the explanation involving the shift from a paratactic to a hypotactic structure cannot straightforwardly be endorsed because the semantics of the embedded subjunctive clause cannot be related to the values of the subjunctive in the main clause. Pinkster cites as one example the temporal clauses introduced by *cum*. Bybee et al. (1994, p. 222) speculate that the substitution of the indicative by the subjunctive in new contexts is a result of a semantic emancipation and reanalysis, whereby the subjunctive loses its original meanings and comes to be more generally understood as an element marking non-assertive embedded contexts.

However, along with Pinkster, other Latinists observe that mood shifts, and more generally, shifts in the frequency of use of entire structures cannot be easily explained by resorting to one general principle. Thus, Grandgent (1907, p. 51) remarks that from Early to Late Latin there is a loss of systematicity in the use of the subjunctive:

At the end of the Vulgar Latin period it was probably used, in popular speech, very much as it is used in the Romance languages. Late writers, while trying to follow the traditional practice, were less logical and evidently less spontaneous than Classic authors in their employment of the subjunctive. [...]

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

late writers often put the subjunctive where Classic authors would have put the indicative (pp. 52-53).

Moreover, the subjunctive came to fill the morphological system of the imperative. Because the latter was restricted to the second person singular and plural of the present, the subjunctive provided the third-person and the first-person forms. While in some cases the subjunctive replaces the indicative or the imperative, largely reverse developments take place: the subjunctive comes to be limited to a smaller number of functions and it is to a great extent supplanted by the indicative and the infinitive. The infinitive clause takes on new semantic domains and comes to be used instead of the subjunctive *ut*-clauses. In particular, Grandgent (1907, p. 53) details the use of modal verb *debeo* “must” with the infinitive, e.g. *debeant accipi = accipiantur* “they must receive”, or the infinitive after *facio* “make/do” in lieu of a subjunctive (see (221)) and the use of the infinitive replacing the embedded indirect interrogative clause where the subjunctive occurs (see sentence (222) from Grandgent [1907, p. 51]). Also, he talks about the internal restructuring of the subjunctive paradigm in Vulgar Latin with certain tenses falling into disuse in favor of others:

The imperfect subjunctive gradually gave way to the pluperfect: this use is common in the *Bellum Africanum* {Lat. Spr. 489}; cf. Sittl 133-134. It apparently began with *debuisset, potuisset, voluisset*, used freely for the imperfect by Gregory the Great (*Sepulcri* 226) and others, and with perfect infinitives

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

like *tacuisse* for *tacere* (Lat. Spr. 489: examples from the 4th century). The imperfect subjunctive ultimately went out of use, except in Sardinia.

(221) *ecce Pater fecit Filium nasci de vergine*

behold, Father make.3.SG.PRF.IND Son originate.INF from virgin
“behold, the Father had the Son be born out of a virgin” (Regnier
27-28 in Grandgent, 1907, p. 53)

(222) *non habent unde reddere tibi*

NEG have.3.PL.PRS.IND ADV render you.DAT
“They do not have anything to render to you.” (from Grandgent,
1907, p. 51)

4.5 Conclusion

The quantitative analyses conducted by Pinkster (2015) on the Latin sample available in the LASLA corpus and the various shifts relayed among others by Grandgent (1907), Ledgeway (2012) and Magni (2010) reveal two important facts. The first is that the subjunctive was largely the mood of subordination even before the emergence and rise of new mood and tense forms (e.g. the conditional and the future); also, even before the more extensive use of modal verbs to encode modalized meanings in the main clause in the Romance vernaculars.

Still in Classical Latin, the subjunctive was the most frequent mood after the indicative. In the Romance languages a quick glance at the absolute frequencies of the moods in the language samples collected in

4 A diachronic analysis of the subjunctive mood from Latin to the Romance languages

the various corpora shows that the subjunctive has been surpassed by other moods. In particular, in Italian the C-ORAL-ROM corpus reveals that, after the indicative, the infinitive has become far more common followed by the conditional. The subjunctive appears to have similar rates as the imperatives.

A second most important observation concerns the semantic relationship between the subjunctive and the indicative in Latin. The two moods did not occupy neatly divided positions in the semantic space clearly identifiable in the factual/ non-factual or counterfactual distinction. Morphological variation also occurred in certain constructions in Latin. Recall in this respect the alternation between the subjunctive and the indicative in indirect questions reported in Magni (2010, p. 247) and shown in sentence (220). Additionally, the functions of the subjunctive were so variable that Roman grammarians explicitly rejected a clear semantic qualification of the subjunctive and, as Pinkster (2015) aptly suggests “[f]or language users as well, especially language learners, it must have been complicated to discover a rationale behind the various uses of the subjunctive in subordinate clauses” (p. 619). This conclusion is particularly relevant for the purposes of the present dissertation as it seeks to investigate how speakers understand and use the Italian subjunctive and more generally how morphological rules are discovered and applied to new language material.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

In the previous chapter, I reviewed the accounts of the subjunctive in Latin produced by the ancient Roman Grammarians as well as modern and contemporary Latinists and linguists. Also, a contrastive analysis between the functions of the subjunctive in Latin and in some of the Romance languages was reported. The comparison revealed that overall, in unembedded utterances the subjunctive has fewer domains of application in the daughter languages French, Spanish and Italian than those that were possible in their common ancestor. Nowadays, some of the occurrences in the main clause are part of completely fossilized or more rigid phrases. Additionally, in the Romance languages, the subjunctive has largely been replaced by other mood forms in the main clause, forms such as the conditional or the modern synthetic future which arose in the development from Late Vulgar Latin to Common Romance. These observations might generate the impression that the mood was fully functioning as a productive semantic marker in the main clause in the era of Classical Latin. Instead, quantitative analyses of Latin texts from the period 250 BC – ad 450 using descriptive statistics revealed that already in the authors of the Roman Republic, the subjunctive was largely a subordinate clause mood and it was understood as such by the grammarians of the time. They portray it as a mood that has no inherent meaning and that relies on other elements of the syntagmatic context to acquire a particular value. The latter point has been stressed more recently by Pinkster

5 A synchronic analysis of mood variability: the parameters that affect mood choice

(2015) who, after considering the uses of the subjunctive in the main and in the subordinate clause in Classical Latin, concluded that, while they might have originated from a single set of core functions, synchronically there is no common denominator among them and for the speakers and language learners of the time, it might have been impossible to abstract away from this variation to a general or unitary semantic value of the mood.

In the present chapter, I review the corpus and experimental studies on the Romance subjunctive. These studies set the focus on actual usage in speech in contrast to the prescriptive and normative accounts that form the standard language tradition cultivated by institutes of education.

Their methods are also essentially different from the approach adopted by the studies within semantics, modal logic, and generative theory presented in chapter 3. Those analyses largely depended on the linguistic intuition of a single individual or a small set of individuals who consciously evaluated the grammaticality of the mood in a number of sentences. The sentences were rarely taken from naturally occurring text but were constructed ad hoc by the researcher often with the goal of testing the boundaries between grammaticality and ungrammaticality. While this approach may be useful during the exploratory stages in order to identify possible linguistic parameters that impact mood choice or key characteristics of the grammar of the mood (recall in this regard the studies on the relation between the subjunctive and complementizer deletion or the *consecutio temporum*

5 A synchronic analysis of mood variability: the parameters that affect mood choice

in Italian), it ultimately ignores how the speech community at large uses the subjunctive in natural interactions situated in a variety of social settings. The semantic studies presented in chapter 3 have no access into the social and cognitive factors that may impact mood choice. Thus, they lack ecological validity and are unable to shed light on phenomena of which the researcher has no direct or conscious experience. In short, they cannot be claimed to be representative of how the subjunctive is actually used by the speech community at large at a given point in time.

The corpus studies reviewed will be supplemented by a corpus analysis of the subjunctive in Italian. The aim of the study is to examine to what extent the occurrence of subjunctive forms is constrained by co-occurring lexical elements. In particular, I investigate the influence of the governor verb (matrix verb) on the subjunctive/indicative distribution and establish whether the lexical semantics of the governor can account for the mood distribution as suggested by semantic-based accounts (Farkas 1992b; Giorgi & Pianesi, 1997 among others: see again chapter 3).

Moreover, an analysis of the productivity of the subjunctive mood in speech is conducted. I examine the frequency distribution of subjunctive forms by verb types and verb tokens and present a simple measure of lexical diversity (type-token ratio). The analysis additionally focuses on the distribution of subjunctives among high-frequency verbs. These properties of the mood are contrasted with those of the infinitive. The latter, being a productive morphological system,

5 A synchronic analysis of mood variability: the parameters that affect mood choice

serves to establish a baseline in order to interpret the data and the patterns that emerge from the analysis. The observations gathered through the corpus study will serve as the basis for the experiment on the factors that influence morphological productivity presented in chapter 6.

5.1 Theoretical background

In adopting an empirical framework, it is also my goal to review and critically evaluate the fundamental assumptions and evidence produced by empirical studies on the Romance subjunctive. In this process, particular attention is directed to the theory of grammaticalization and to the studies on the relationship between frequency of occurrence of linguistic elements and their cognitive representation.

In the next sections, I will introduce the theories developed by the studies on language variation and language change and the psycholinguistic theories on the role of frequency in language. I will then turn to the premises and methods of corpus linguistics and present the corpus resources used for the corpus study.

5.1.1 Grammaticalization

The studies conducted within two linguistic traditions, Indo-European historical linguistics and language typology, (Humboldt, 1822; Bopp & Windischmann, 1816; Meillet, 1912; Hodge, 1970; Givón, 1971; Kuryłowicz, 1965) have closely examined how languages change over time. They developed the concept of *grammaticalization* which

5 A synchronic analysis of mood variability: the parameters that affect mood choice

originally described the development of grammatical material out of lexemes and the evolution of affixes (bound morphemes) from free morphemes. One of the first definitions is provided by Kuryłowicz (1965):

“Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one (p. 52).”

Already in this early characterization, Kuryłowicz dwells on the distributional properties of the morphemes that become grammaticalized. As Lehmann (2015, p. 7) clarifies, “increase of the range” refers to a wider scope of application of the morpheme, which can be applied to increasingly more lexemes. This measure is also known as type frequency (Bybee & Thompson, 1997). Lehmann (2015) essentially adopts Kuryłowicz’s definition highlighting that grammaticalization is a “gradual shift to a new function” (p. 14). The distinction between lexical and grammatical is not a binary one. They are rather two ends of a continuum. Most importantly, grammaticalization does not end with the shift from the lexical to the grammatical or with the shift from the analytic to the synthetic, although these transitions may be the most salient ones (p. 14). It is rather an “open-ended process” that operates in the same way before and after those turns. Figure 5.1 adapted from Lehmann (2015, p. 15) and based on the original illustration from Givón and Whitaker (1979, p. 209), highlights the main phases of grammaticalization.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Phase	syntacticization	>	morphologization	>	demorphemicization	>	loss	
Level	discourse		syntax		morphology		morphophonemics	zero

Figure 5.1: The phases of grammaticalization

Adapted from *Thoughts on grammaticalization* by Lehmann (2015, p. 15). Copyright 2015 by Language Science Press.

As Figure 5.1 illustrates, grammaticalization may end with the complete loss of grammatical material. However, as Lehman (2015, p. 15) notes, not all linguistic forms experience the entire cycle of grammaticalization. In fact, it is not a necessary condition that every form go through all the stages. Moreover, the form can enter the grammaticalization process at any stage. The final outcome is not predetermined either.. All these observations are particularly relevant for the object of study of the present dissertation. The subjunctive is a morphological paradigm made up of synthetic inflectional affixes. In the figure above, it can be placed on the second to last phase of the grammaticalization process, in the morphophomenics category. Nonetheless, it can still be subject to the mechanisms of grammaticalization. The most relevant to the current study are discussed in the following.

Phonological attrition (Lehmann, 2015, p. 134) or erosion (Heine & Reh, 1984) is understood as the loss of phonological material. The forms become shorter and progressively lose phonetic features. The loss of phonetic material naturally affects morphemes as well. Morphological simplification, the loss of distinctive features within a paradigm, or morphological change can also take place due to

5 A synchronic analysis of mood variability: the parameters that affect mood choice

analogical pressure, whereby a morpheme is substituted by another that has the same or a similar function. Usually the new morpheme substituting the older variant has a larger distribution, i.e. a higher type frequency. This means that it occurs together with more lexemes, a factor that may increase its absolute frequency of occurrence as well (for a more in-depth discussion on the role of frequency in language change see section 5.1.2). A famous example of analogical simplification in the history of Latin is the substitution of the phoneme [s] with [r] in the nominative form *honos* → *honor* “honor” with the effect of eliminating the stem final contrast between [s] and [r] in *honos* ~ *honoris* “honor.NOM/GEN” → *honor* ~ *honoris*. This change takes place by analogical extension of the morphological pattern of other Latin lexemes like *decor* ~ *decoris* “grace.NOM/GEN”, *soror* ~ *sororis* “sister.NOM/GEN” or *labor* ~ *laboris* “work.NOM/GEN” (Hock, 1991, pp. 179-190; Barr, 1995, pp. 509-544; Kiparsky, 1978; Albright, 2008, p. 7; Albright, 2005).

In chapter 2, I discussed the phenomenon of morphological leveling affecting the present tense of the Italian subjunctive. In natural discourse inflectional affixes of the 2nd or 3rd conjugation are sometimes substituted with the affixes of the 1st conjugation. This extension of the paradigm of the 1st conjugation to verbs of other conjugation classes generates forms that are considered errors in the standard language: for instance *venghino* instead of the correct form *vengano* “come.3.PL.PRS.SUBJ” for verb *venire* belonging to the 3rd conjugation, or *sappi* instead of *sappia*-1/2/3.SG.PRS.SUBJ “for *sapere* “know” (see

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Table 5.1 for the inflectional paradigm of the subjunctive present according to conjugation).

Table 5.1
Inflectional paradigm of the Italian subjunctive present according to conjugation classes

1 st conjugation <i>pensare</i> “think”	2 nd conjugation <i>fare</i> “feign”	3 rd conjugation <i>coprire</i> “cover”
<i>pens-i</i>	<i>fa-a</i>	<i>copr-a</i>
<i>pens-i</i>	<i>fa-a</i>	<i>copr-a</i>
<i>pens-i</i>	<i>fa-a</i>	<i>copr-a</i>
<i>pens-iamo</i>	<i>fa-iamo</i>	<i>copr-iamo</i>
<i>pens-iate</i>	<i>fa-iate</i>	<i>copr-iate</i>
<i>pens-ino</i>	<i>fa-ano</i>	<i>copr-ano</i>

The 1st conjugation is the largest in Italian, encompassing most verbs. Based on this fact, one might assume that the subjunctive affixes of the 1st conjugation have the widest application and should exhibit the highest type frequency in discourse (for a detailed discussion of the statistics of occurrence of the subjunctives of each conjugation see section 5.3.3.2). However, in chapter 2, I also reported the occurrence of errors in natural discourse due to the extension of the subjunctive paradigm of the 2nd or 3rd conjugation to verbs of the 1st, which shows the opposite trend. I take these phenomena as an indication of the high degree of uncertainty of the speakers concerning the morphology of the

5 A synchronic analysis of mood variability: the parameters that affect mood choice

subjunctive. Following Barr (1995), Garrett (2008), and Albright (2008) among others, I hypothesize that these analogical extensions are generated by the lack of sufficient input based on which the speakers successfully produce appropriate forms. The mood has a precarious status in the mind of the speakers due to its overall very low frequency in natural discourse. Recall that it is mainly the mood of subordination where it is often replaced by other morphology (see chapter 5 for an extensive analysis of the occurrence of the subjunctive in speech). As it is already very rare in speech, it is impossible to regularly hear all the forms of the paradigm, all the more because they vary according to verb person, tense and conjugation, thereby generating a highly complex inflectional system.

Analogy takes place as a form of inferential process to derive the necessary form when there is not enough evidence of a form or a paradigm in the linguistic input that the speaker or learner receives. Speakers “need to make inferences about the phonological and morphological properties of words based on incomplete information” (Albright, 2008, p. 7) and will select the form which they can confidently generate, of which they have enough evidence (see chapter 6 for a study on the mechanisms of rule generalization). Thus, Albright emphasizes that

[...] nothing in this system requires that overregularization/leveling must take place; as long as learners have sufficient access to input data, there is always the potential

5 A synchronic analysis of mood variability: the parameters that affect mood choice

to learn and maintain irregularity in derived (=non-basic) forms. [...] morphological change is driven diachronically by inaccurate or incomplete transmission of the full set of inherited forms (p. 8).

This view of analogy connects morphological variation and change to morphological learning and to the cognitive representation of morphological knowledge. The principles that underlie language learning like analogical inference ultimately impact the forms that speakers produce and transmit.

It is important to note that this account of analogy differs from the traditional understanding developed in functional theories of the late 20th century (Haiman, 1985, 2010, p. 159; Langacker, 1977) or in Optimality Theory (McCarthy, 2002; Kiparsky, 1968; King, 1969, chapter 5.3). Those studies assumed that analogy is as a realization of the “iconicity” principle, a result of the cognitive drive toward order, transparency and simplicity, toward achieving a “one form = one meaning” correspondence and toward eliminating ambiguity and exceptions with the benefit of obtaining maximum communicative clarity. This explanation comes from the observation that analogical change largely involves the substitution of an exceptional or irregular form with a regular one. It is usually based on the extension of the least marked form (Jakobson, 1939; Greenberg, 1966; Bybee & Brewer, 1980; Bybee, 1985) or the least suffixed member of the paradigm (Mańczak, 1957; Bybee, 1985) to the marked form (unidirectionality of analogical change). For a more extensive review see Albright (2008, p.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

4). However, Haiman (2010) gives ample evidence of cases where analogical change does not result in a perfect form/meaning correspondence, where “[...] paradigm coherence or uniformity is imposed at a cognitive cost: semantic distinctions and clarity are lost for the sake of a pattern that is 'tidy' [...]" (pp. 159-160). Many analogical changes do not bring about simplification and it is not uncommon for the marked or suffixed form to become the base of the analogical extension (see Vennemann [1972] and Albright [2008] for evidence coming from Yiddish and German among other languages).

In the present study, analogical simplification is not taken to be a necessary outcome. It is a process that sets in when there is incomplete knowledge and has to do with the specific mechanisms that allow for that incomplete knowledge to be filled.

Desemanticization (also known as semantic bleaching or semantic attrition) characterizes the progressive loss of meaning, lexical or grammatical, in the course of grammaticalization. In the context of morphology, “[it] occurs as a morpheme loses its intention: From describing a narrow set of ideas, it comes to describe an ever broader range of them, and eventually may lose its meaning altogether" (Haiman, 1991, p. 154). I put it here together with *coalescence*, another parameter of grammaticalization. It describes an increase in bondedness between one form and other linguistic elements of the syntagmatic chain: “a grammaticalized sign only becomes meaningful when it occurs in the company of other elements” (Dragosits, 2017, p. 9).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

The debate on whether the subjunctive can be considered a semantic marker or whether it has lost its semantic value(s) altogether has been extensively reviewed for both Latin and the Romance languages in the previous chapters. In the corpus study in section 5.3, I investigate whether the subjunctive coalesces with other elements of the sentence, like the governor, thereby forming specific collocations which cannot be motivated by underlying semantic factors.

Obligatorication is a neologism introduced by Lehmann (2015) to describe the process by which a grammatical element progressively becomes obligatory. As a result, the language user can no longer choose whether to express the grammatical category and its values or not. Lehmann gives as an example the development of definite and indefinite articles from Latin to the Romance languages. In Latin, the realization of the determiners *ille* (“that”) or *unus* (“a”) was not obligatory. In the Romance languages, however, “it is in most contexts impossible to use a singular noun – and, with local variation, also a plural noun – without an article, that is, without specifying the category of definiteness.” (p. 149). In the context of the subjunctive, instances of elements becoming obligatory in the syntax of subjunctive phrases will be discussed.

5.1.2 Frequency effects in language

The role of frequency has been widely researched in a variety of language processes. Linguistic productivity, language change, language learning and language processing are all shaped by frequency

5 A synchronic analysis of mood variability: the parameters that affect mood choice

of usage. Evidence from these domains of study suggests that the human brain is sensitive to the statistical properties of the linguistic input. This statistical knowledge is used by the learner/speaker to make predictions about incoming input and about the appropriate linguistic output to produce:

The more times we experience something, the stronger our memory for it, and the more fluently it is accessed. [...] The more times we experience conjunctions of features, the more they become associated in our minds and the more these subsequently affect perception and categorization; so a stimulus becomes associated with a context and we become more likely to perceive it in that context (Ellis, 2012, p. 7).

While the statement is per se uncontroversial, there is nowadays still a very lively debate concerning which frequency measures are actually used by the brain and how they interact with the different levels of linguistic structure in the various processes connected to language comprehension and language production (Yang, 2004, 2015; Baayen, 2010; Amenta & Crepaldi, 2012; Kapatsinski, 2012). There is even less consensus with regard to how frequency effects in language are to be interpreted in more general theories, those that speculate about how linguistic knowledge is generated and examine the relationship between language-specific processes and general cognition (Chomsky, 1981; Bybee & McClelland, 2005; Ramscar & Gitcho, 2007; Yang, 2010; Kidd, 2012).

5 *A synchronic analysis of mood variability: the parameters that affect mood choice*

In the next sections I will present the frequency measures that have been considered in studies of language productivity and language change. An in-depth discussion of the role of the various frequency measures in some of the most important models of morphological productivity is tackled in chapter 6.

5.1.2.1 Frequency and productivity

At a broad level, two frequency measures, type and token frequency, are consistently mentioned in the studies exploring the role of frequency on language productivity and language obsolescence. Token frequency counts how often a specific form appears in the language input. It is usually extracted from a corpus taken to be representative of the language input of a given speaker. For instance the word form *forsaken* (past participle of *forsake*) counts an absolute token frequency of 1191 in the COCA corpus (Davies, 2008), 1.09 occurrences per million words. Its token frequency differs from the one exhibited by the lexically related form *forsook* which in comparison registers only 132 occurrences overall, 0.12 per million. Type frequency refers to the number of distinct items (types) that take a certain pattern or that fill a slot in a relevant structure (Bybee & Thompson, 1997, p. 378; Ellis, 2012, p. 12; Pfänder & Behrens, 2016; p. 9). At the morphological level, the English past tense suffix *-ed* has a high type frequency because it is encountered with many different lexical items in discourse. Indeed, most verbs take this suffix to generate the past tense. Per contra, the past tense pattern involving the ablaut [ɪ]→[æ] of the verb stem has

5 A synchronic analysis of mood variability: the parameters that affect mood choice

comparably a very low type frequency, only a small set of verbs that still occur in speech participate in this morphological change, e.g. *swim* → *swam*, *drink* → *drank*, *ring* → *rang*. Type frequency is correlated with morphological productivity: the pattern that is used with most word types is also almost always the one employed to derive new word forms (see section 4.4 in Yang [2016] for a discussion on the productivity of the *-s* suffix in German noun plurals, a pattern with relatively low type frequency). Beyond the mere correlation, however, type frequency is considered to play a fundamental role in the acquisition of the productive morphological pattern (Bybee, 1995). The child has to hear a given morphological change sufficient times with sufficiently diverse words, in order for her to discover a rule to apply productively for the formation of morphologically complex items:

No matter how many times the child hears a single verb type that takes the *-d* suffix (e.g., *talked*), it would be folly to draw the sweeping conclusion that “*-d*” can attach to all verbs. A productive rule must be supported by a sufficiently large number of distinct types such that its open-endedness can be justified. Thus, rules do not become productive overnight and can only emerge after the accumulation of evidence (Yang, 2016, pp. 67-68).

A large and varied domain of application is necessary in order to generalize a pattern to other words independently of their semantic, morphological, phonological or prosodic structure: if the members of a

5 A synchronic analysis of mood variability: the parameters that affect mood choice

set to which a morphological pattern applies are very diverse, no distinctive similarities (subregularities) among them emerge, that provide the learner with a cue that the rule is restricted to a specific environment. For a more in-depth analysis of how different frequency measures impact productivity see chapter 6.

5.1.2.2 Frequency and obsolescence

A morphological system becomes obsolete and is no longer generalized to new words when there is not enough evidence of it in the language input that language learners receive relative to other morphological rules that are available to encode the same function. It suffers from the competition of other patterns that have a larger domain of application (high type frequency) and are thus extended to increasingly more words. While exhibiting a low type frequency, the obsolescing system is usually found to be preserved in tokens that are highly frequent in discourse. The most cited example undoubtedly concerns the English past tense system. In modern English, irregular past tense forms mostly belong to verb forms that exhibit high token frequency, among them auxiliaries or the modals (e.g. *was, had, did, went, said, could, might* etc.). Bybee & Slobin (1982) noted that twenty-two out of the thirty most frequent English verbs reported in Kučera and Francis (1967) are in fact irregular. In a corpus study on the regularization of English verbs from Old English to Modern English over a time span of 1,200 years, Lieberman et al. (2007), focusing on a set of 177 verbs, found the rate of regularization to be inversely correlated with verb lexical frequency:

5 A synchronic analysis of mood variability: the parameters that affect mood choice

more frequent verbs regularize at a slower pace. In fact, “[they] regularize at a rate that is inversely proportional to the square root of their usage frequency” (p. 714).

Thus, while high type frequency is necessary for the analogical extension of a pattern thereby contributing to its productivity, word forms with very high token frequency tend to resist analogical leveling.

5.1.2.3 Frequency and the mental representation of complex lexical units

The conserving effect of token frequency (Bybee & Thompson, 1997, pp. 380-381) has been addressed by psycholinguistic theories of lexical processing and lexical access like exemplar-based models (Levelt et al., 1999; Pierrehumbert, 2002; Levelt, 2002; Baayen, 2003a), the network model (Bybee, 1988, 1995), connectionist models (Daugherty & Seidenberg, 1994; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1986; McClelland & Patterson, 2002) and connectionist distributed models (Gaskell & Marslen-Wilson, 1997; Devlin et al., 1998; Seidenberg & Gonnerman, 2000) among others⁸. These models posit that frequency of occurrence affects the mental representation of complex word forms and of entire multi-word phrases. Repeated exposure to a particular word form strengthens its mental representation or its resting activation level understood as a determinant of its position of retrieval and leads to its stronger entrenchment in memory

⁸ See also the Word and Paradigm Model in Matthews (1974) and Blevins (2003, 2006).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

(Blumenthal, 2017; Langacker, 2017). When encountered very frequently, a complex morphological form leaves a trace in long-term memory and it can be accessed directly without parsing its constituent parts first. In that case, the internal structure will not be processed or it will be processed only shortly after (Schreuder & Baayen, 1995; Baayen & Dijkstra, 1997). This mechanism is also called *chunking* and it is assumed to operate at the syntactic level as well. Frequently co-occurring words become part of larger phrasal units which are routinized and are perceived as chunks, they are *lexicalized* (Sosa & MacFarlane, 2002; Kapatsinski & Radicke, 2009): the constituent elements lose independence and the structure becomes fixed. Among others, studies conducted by Arnon & Snider (2010) provide evidence that four-word phrases are processed faster as a function of their frequency and may be stored in lexical memory.

As the internal structure of highly frequent lexical items or phrases is not accessed, it resists the pressure to be substituted by more productive patterns, like the shift to the regular English past tense suffix *-ed* already discussed (Bybee & Thompson, 1997).

It is important to keep in mind that phenomena concerning language change like those discussed here are usually determined in diachronic corpus studies where very broad frequency measures are controlled for, like type and token frequency addressed above. However, they are explained by resorting to psycholinguistic theories of how complex lexical items are accessed in on-line processing and how they are represented in the brain.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

The experimental studies that are designed to directly test these processes, like online visual or auditory word recognition (with or without priming) have indeed explored the role of lexical frequency on word priming and word recognition time. In particular, the lexical decision experiments conducted on morphologically complex words, like inflected or derived words, have evaluated a number of frequency-related predictors: stem frequency, affix frequency and whole-word frequency. In these studies, “the existence of a frequency effect for a given unit (syllable, simple word, complex word, phrase) is often interpreted as empirical evidence for the existence of cognitive representations for such units” (Baayen, 2010, p. 2) which become relevant processing units. According to the review by Amenta & Crepaldi (2012), across visual word recognition experiments, the most robust results confirmed the effects of both stem frequency and whole-word frequency (also called surface frequency) on the speed of word recognition (Taft, 1979; Taft & Ardasinski, 2006; Burani et al., 1984; Cole et al., 1989; Baayen & Dijkstra, 1997; New et al., 2004; Baayen, 2003b; Baayen et al., 2006; Baayen et al., 2007). Interestingly, stem frequency interacted with whole-word frequency (Caramazza et al., 1988; Beauvillain, 1996, Baayen & Dijkstra, 1997; Schreuder & Baayen, 1997; Alegre & Gordon, 1999; Allen et al., 2003; Kuperman et al., 2010). For instance, in a study by Burani & Thornton (2003) on lexical decision times, “stem frequency had a facilitatory effect for the lowest frequency words, but an inhibitory effect for the highest frequency words” (Amenta & Crepaldi, 2012, p. 2). This interaction

5 A synchronic analysis of mood variability: the parameters that affect mood choice

could indicate that high-frequency lexical items are processed without accessing the stem and affix first and it is taken as evidence for models that posit storage of morphologically complex forms as a result of their frequency.

Finally, experiments on morphological priming provide evidence that words are recognized faster when the morphologically related prime is infrequent. High-frequency primes that have the same morphological relation with the target word as the low-frequency ones do not yield a comparable processing advantage (Raveh, 2002; Meunier & Segui, 1999; Giraudo & Grainger, 2000). The results support the hypothesis that the internal structure of high-frequency primes is not processed. However, Amenta & Crepaldi (2012) note that especially in masked priming, which should reflect early stages of lexical access these effects are still disputed (McCormick, 2009).

Overall, surface frequency emerges as an important parameter in the processing of morphologically complex word forms. However, Baayen (2010) for one argues that its role is over-estimated due to the type of statistical analyses conducted which mainly involve multivariate linear regressions. According to his study, lexical frequency actually masks other frequency-related variables which are collinear and predict lexical frequency. Some of the variables are: *contextual diversity* (or *text dispersion*), a measure that counts the number of different texts a word appears in or the *microcontext of a word*, i.e. the close syntactic context

5 A synchronic analysis of mood variability: the parameters that affect mood choice

in which a target word appears. This usually comprises four or five words before and after the target word.

There are, however, also morphological predictors that reflect the influence of paradigmatic relationships among words on morphological processing. Some of the predictors investigated are *inflectional entropy*, a measure of the frequency distribution of inflectional endings of an inflectional class⁹ (del Prado Martín, 2004; del Prado Martín & Baayen, 2005; Milin, Kuperman et al., 2009), *paradigm entropy* capturing the frequency distribution of the affixes of a paradigm or *morphological family size* (Schreuder & Baayen, 1997; Bertram et al., 2000) defined as “the type count of lemmas (both derived words and compounds) that share a given base word (a stem) as a constituent” (Traficante & Burani, 2003, p. 46).

They highlight the role of contextual learning: “frequency reflects a wide range of lexical distributional properties that are all co-determining learning” (Baayen, 2010, p. 2). For instance, McDonald and Shillcock (2001) found that the more fixed the syntactic environment in which a word occurs (the more the word is part of a collocation), the longer the time required to recognize it will be. This measure is captured in Baayen in the microcontext of a word.

⁹ Inflectional class is defined as the “[group] of lexemes that use the same set of exponents in their inflectional paradigms” (Milin, Kuperman et al., 2009, p. 1). It is different from the inflectional paradigm understood as “the set of inflected variants of a given lexeme” (Milin, Kuperman et al., 2009, p. 1).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Thus, some processes may enhance the speed of lexical processing while others may inhibit it in an interactive manner (Stockall et al., 2004; Baayen et al., 2011; Ramscar & Baayen, 2013; Milin et al., 2017; Milin et al., 2018). In the case of morphological processing, they may tip the scales either in favor of whole-word access or in favor of decomposition into sublexical structures like constituent morphemes (see for one the role of inflectional entropy and paradigm entropy in visual word processing investigated in del Prado Martín et al. [2004], Ferro et al. [2018, p. 322], Milin, Kuperman et al. [2009], Milin, Đurđević and del Prado Martín [2009], Luce [1986], Luce & Pisoni [1998] among others).

Most importantly, “learning is driven by both positive and negative evidence” (Milin et al., 2018, p. 12). More advanced models should be able to capture the negative evidence as well in terms of competition.

5.1.2.4 Frequency and the erosion of linguistic structure

Frequency of use affects the acoustic articulation of words and multi-word sequences. High-frequency words are overall produced faster than low-frequency words (Jurafsky et al., 1998; Bell et al., 1999; Jurafsky et al., 2001; Jurafsky et al., 2002). Also, they tend to have reduced consonants or vowels (Hooper, 1976). Thus, in early studies on English, final past tense morphemes /t/ or /d/ were found to be longer in low-frequency verbs than in high-frequency verbs (Fidelholtz, 1975). In addition to facilitating consonant and vowel reduction, lexical frequency was found to increase the probability of phonetic deletion.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Bybee (2002c), based on a previous study (Bybee, 2000) reports that “[r]egular past tense /t/ and /d/ are more likely to delete in high-frequency verbs than in low-frequency verbs.” (p. 265)

There is by now evidence that frequency does not operate exclusively at the lexical level. Frequency relations between words are equally important. Jurafsky reports that word forms tend to be reduced when they have a higher probability of appearing in a specific lexical context, i.e. when they are predictable in a given context. More specifically, he measures the so-called *transitional probability*, i.e. the conditional probability of a word given the previous or the following word (Bush, 1999; Saffran et al., 1996). The transitional probability is estimated based on the frequency of co-occurrence of two words in a large corpus: the number of times in which two words occur together is divided by the number of times the first word occurs (forward transitional probability) or by the number of times the second word occurs (backward transitional probability)

The results of their analyses indicate that the predictability of the word based on the following word affects its duration: a word is likely to be reduced if it is highly predictable given the subsequent word. Thus, first words of frequent word pairs like “east coast”, “need trimming”, “last resort” tend to be shorter than first words in infrequent word pairs (a summary of the results is reported in Bell et al. [2009]; see Bybee [2002b] and Schäfer [2014, pp. 9-11] for additional reviews).

Frequency of co-occurrence is reported to blur the boundaries between independent segments at different structural levels, phonemic,

5 A synchronic analysis of mood variability: the parameters that affect mood choice

morphemic or even at the word level and it is considered a driving force for reduction in the course of grammaticalization. The frequent English semi-modals *going to*, *(have) got to*, and *want to* (Biber et al., 1999; Hopper & Traugott, 2003) are now often reduced to *gonna*, *gotta* and *wanna*, in which the infinitive marker *to* and the preceding modal verb, formerly independent entities, have fused together. The fusion and reduction of the forms is often accompanied by a functional reanalysis. In the case of the afore-mentioned semi-modals, the contracted variants have come to express different aspects of modality than the source forms and are used in different syntactic positions (see Lorenz [2013] for evidence and discussion on the role of frequency in the semantic emancipation of the contracted semi-modals). Similarly, the entire phrase *I am going to* is now reduced to *I'mna* ['aimənə] due to its high token frequency (Bybee & Thompson, 1997, pp. 379).

Different explanations have been proposed for the effect of word frequency and word predictability on their articulatory realization (for a comprehensive review see Bell et al. [2009]).

Bybee & Hopper (2001) and Bybee & McClelland (2005) suggest that frequency directly facilitates articulation: through regularly repeated production the (co-)articulation of word forms and word sequences is highly practiced and becomes automatized. However, according to Bell et al. (2009, p. 95), this account fails to explain the range of results coming from the studies on word predictability or those on syntactic predictability (Gahl & Garnsey, 2004; Fisher et al., 2006). Other explanations draw on the theoretical work which posits a

5 A synchronic analysis of mood variability: the parameters that affect mood choice

fundamental relationship between frequency and the cognitive representation of word forms, like the exemplar model (Levelt, 2002; Pierrehumbert, 2002) or the models of lexical access presented in sections 5.1.2.3 and 5.1.2.4. Altogether, models based on lexical access make the following prediction for speech production: lexical items that are retrieved faster due to their high activation level in the brain are also produced faster. This hypothesis presupposes the existence of a coordination mechanism which links a word's lexical activation to its phonological encoding. See Bell et al. (2009) for an outline of this mechanism. Combining cognitive and articulatory processes, which have traditionally been studied separately, remains challenging (Mousikou & Rastle, 2015, p. 6).

5.1.3 Conclusion

In section 5.1.2, I reviewed the literature on frequency effects in language. In particular, I discussed the studies investigating how usage statistics shape the mental representation of complex lexical units, reporting the experimental evidence gathered by the studies on language comprehension and production.

It was also discussed how frequency indirectly affects language variation and language change. In particular, it was shown how changes concerning reduction, fusion and loss usually affect high-frequency words or high-frequency word sequences first and later spread to other lexical elements. On the other hand, changes due to the extension of a productive pattern first spread to infrequent words while high-

5 A synchronic analysis of mood variability: the parameters that affect mood choice

frequency items tend to resist it (Bybee, 2002b; Hooper, 1976, pp. 269-271; Lieberman et al., 2007). At a given moment in time, this phenomenon generates variability in the language system. In section 5.3 I explore how these processes affect the occurrence of the subjunctive mood in Italian. In particular, I assess to what extent the subjunctive is lexically conditioned by the governor verb and by high-frequency embedded verbs.

5.1.4 Language variation and change

Sociolinguistics has given particular relevance to the study of language variation in everyday speech. Loosely defined, variation is the alternation of two or more forms used to express the same thing (Labov, 1972, p. 8). When the forms alternate or compete in the same discursive contexts to encode the very same or similar discursive functions (Sankoff & Thibault, 1981, p. 207), they are considered variants of the same linguistic variable. A variable can be phonological, morphological, morphosyntactic, lexical or discursive (Tagliamonte, 2011, pp. 2-8). In short, variable linguistic behavior is found at all levels of grammar. Moreover, sociolinguists observe that variation is pervasive in or rather “inherent to” the language variety of every community, present or past, (Labov, 1969, p. 728). Crucially, from the empirical studies conducted in the late sixties, variation appears not to be random, nor completely determined (Lorenz, 2013, p. 4). It is rather systematic. The notion of *structured* or *orderly heterogeneity* is invoked to characterize this fact. The choice of one variant over the

5 A synchronic analysis of mood variability: the parameters that affect mood choice

other competitors cannot be understood by positing a categorical rule. In order to map this type of variation, one must investigate the factors that influence it in a probabilistic manner. Variation must therefore be quantitatively modeled (Labov, 1963, 1969) using statistical methods (Sankoff, 1988, p. 151) that weigh the significance of each factor and its effect size. The factors are either language-internal or language-external. Language-external factors are social in nature, they pertain to the social identity of the speaker within the speech community.

According to Labov (1972), “[t]he speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms [...]” (p. 248). Within this larger social environment, the individual speakers are characterized by a number of dimensions like age, sex, gender, socioeconomic class, occupation, ethnicity, level of education and also finer-grained identity traits that become relevant in their daily interactions (see Eckert [2006, 2012] for the relevant factors considered in the “third wave” of sociolinguistics). All these variables may influence language behavior and, in particular, impact the choice of a linguistic variant. Additionally, there are language-internal or cognitive factors that pertain to the speakers’ knowledge of the grammar. They are usually established in terms of correlation between the variable that is the object of study and a number of linguistic elements within the discourse context.

The study of variation has been particularly relevant for the study of language change and in this respect it supplements the insights of

5 A synchronic analysis of mood variability: the parameters that affect mood choice

historical linguistics and language typology providing the missing link between the synchronic study of language viewed as the abstract knowledge of a formal system shared by the speech community at a given point in time, and the diachronic perspective that highlights the dynamic nature of language. It also emphasizes the social dimension of change. Herzog et al. (1968) posit that synchronic variation is a fundamental ingredient for language change: “Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity.” (p. 188). Change starts from an initial stage of variability in which two variants coexist, although they do not have a comparable distribution across the language community. Change is understood as diffusion or propagation of the new variant from a restricted group of speakers to the larger speech community over a substantial period of time. While at the beginning, the new variant is highly marked, i.e. it may be perceived as incorrect, non-standard or a distinct characteristic of the talk of a particular social group, it is later integrated into the speech of the larger community. It is thus learnt as the default by subsequent generations of speakers. It loses its *markedness*, which shifts to the older form gradually being replaced: “At the latest stages of change, there may be little correlation [between the new variant and] social factors”. (Herzog et al., 1968, p. 185)

An important method introduced by sociolinguistics is the study of change in apparent time (Bailey et al., 1991). The linguistic production of different age groups is taken as representative of different temporal stages of the language. It reflects the time in which each generation

5 A synchronic analysis of mood variability: the parameters that affect mood choice

acquired the language variety. Generational differences in the speech forms employed are thus interpreted as diachronic changes in progress in the language. More specifically, the frequency with which a form is used across different generations is an important measure for determining a shift (Tagliamonte, 2011, p. 43). Another more traditional method concerns the analysis of samples of language coming from different ages with the aim of reconstructing how particular forms developed throughout time. The quality and quantity of the language samples available greatly determines the potential of the analysis. Also, the detail and level of annotation of large collections of language samples (corpora) enables a more comprehensive mapping of different factors including sociolinguistic variables like register, age of speaker, or modality (written or spoken) that may impact language variation. The evolution of linguistic features is reconstructed through statistical analyses that can evaluate both the domains of use of a particular variant and the shifts in frequency of use in each domain across time.

Variationist sociolinguistics provides an important theoretical framework and necessary methods for the synchronic study of the grammar and use of the subjunctive. Following this framework, the subjunctive is understood as a variant of a morphological variable that can take other values in natural discourse such as the indicative or the conditional. All these mood forms are considered variants of the same variable if they appear in the same syntactic, pragmatic and discourse contexts to convey the same functions: no fundamental shifts in meaning arise from the choice of one variant over its competitors. The

5 A synchronic analysis of mood variability: the parameters that affect mood choice

goal of mapping morphological variation comes from the assumption that the alternation is not random but responds to both language-internal and language-external parameters. Modeling those parameters can provide the researcher with an understanding of the grammar of obsolescing paradigms like the subjunctive and give insight into the mechanisms driving morphological productivity and morphological attrition.

Empirical studies like the study of synchronic language variation in speech production make use of both corpus-based and experimental research. The difference between the two methodologies lies within the amount of control that the observer can exercise over the factors which possibly influence the linguistic phenomenon studied (Schäfer, 2014, pp. 30-31). In corpus research, the language sample that constitutes the object of study has been produced spontaneously: “the data are from natural contexts; thus, they make it possible to study register/genre questions that are difficult to study experimentally and come with a higher degree of external validity than some experimental designs” (Gilquin, 2009, p. 9).

However, being natural and spontaneous speech, corpus data varies along a high number of dimensions. The analyst has only limited control over all these variables. Also, annotated corpora of oral language are still relatively small nowadays. In an attempt to control for a high number of factors, the researcher runs the risk of excessively parceling the data: the amount of observations that respond to the values of all the variables considered becomes smaller the more variables

5 A synchronic analysis of mood variability: the parameters that affect mood choice

which are included in the analyses. It becomes impossible to make reliable statistical generalizations over the entire population based on a small set of observations.

Thus, corpus research is especially suited for exploratory research, i.e. research that aims to uncover interesting patterns in the data and possible correlations with a number of variables that enable the researcher to develop relevant hypotheses. Nevertheless, corpus studies have also attempted statistical inference and hypothesis testing (Lüdeling & Kytö, 2008).

On the other hand, ideally, the data gathered through experimental research is carefully controlled for all the relevant variables (Schäfer, 2014, p. 30), both the variables that are explicitly tested and systematically manipulated by the experimenter and the possible confounding variables, i.e. the variables that obscure the impact of the factors under investigation. Consequently, the risk of confounds is minimized. Also, experimental studies are carefully designed in order to elicit enough observations for each value of the variables controlled. In this way, the studies ideally contain enough relevant empirical observations in order to make reliable inferences concerning the linguistic behavior of the entire population considered. Experimental designs, in order to achieve this level of control, do not take place in a natural environment. The input data that the experiment participants receive is carefully constructed and the behavior elicited from the participants, whether linguistic or non-linguistic may not reflect how spontaneous behavior concerning linguistic interaction takes place.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Experiment participants may implement cognitive mechanisms that do not reflect the ones that naturally occur in the process under investigation. The conditions and setting of the experiment are artificial. This must be carefully evaluated by the researcher when analyzing the results of the experiment. According to Gilquin & Gries (2009), the advantages and disadvantages of the two methodologies, corpus studies and experiments, are somewhat complementary and they can be fruitfully combined.

5.2 Corpus-based and experimental research on the subjunctive and on mood variability in the Romance languages: a critical review

In the previous sections, I set out a theoretical framework which includes principles of grammaticalization, as well as sociolinguistic and cognitive mechanisms. They open up new avenues of research into the Romance subjunctive and highlight the role of a wide range of language-internal and language-external factors that may influence the mood variation encountered in the Romance languages. Mapping these variables and their interaction becomes methodologically possible through the implementation of corpus-based and experimental studies.

A number of corpus-based studies have been conducted on the distribution of the subjunctive in spoken Quebec French (Poplack, 1990, 1991; Poplack et al., 2013). Based on her corpus observations, Poplack (1991) notices an *inherent variability* characterizing the use of the subjunctive in speech. The expression is derived from the concept of variation developed in sociolinguistics and describes the alternation

5 A synchronic analysis of mood variability: the parameters that affect mood choice

between the subjunctive and other mood forms in the very same morphosyntactic and lexico-semantic environment with no change in referential meaning. Even if the subjunctive, the indicative and the conditional may have different meanings and functions in certain environments, for instance in unembedded utterances, those differences are neutralized in particular discourse contexts (Sankoff, 1988). Identifying these contexts is to find the *locus of variation* of the morphological variable which can take as its values the subjunctive, the indicative and the conditional. Poplack et al. (2013) note that defining the locus of variation is a challenge in itself. As it was shown in chapter 3, there is little consensus in the literature on the meaning of the subjunctive and on whether the subjunctive/indicative alternation generates a semantic shift or not in the various contexts.

To tackle this issue they adopt an exploratory corpus-based approach:

every instance of subjunctive morphology in the corpus was located and the governor under which it occurred was noted. The set of governors thus identified was taken to be the locus in which subjunctive selection was a possibility. In a second step, all verbs embedded under these governors were exhaustively extracted, regardless of their morphology. By determining the identity of potential subjunctive governors, this procedure not only yields an accountable circumscription of the variable context, it also enables us to assess the propensity of each governor to select for a given morphology, as well as its frequency in running speech (p. 163).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

They acknowledge that this particular procedure may exclude governors which potentially take the subjunctive but appear exclusively with the indicative and the conditional in the collection of texts available in the corpus. However, this method has the benefit of fine-tuning the semantic generalizations about the use of the subjunctive advanced in the literature (p. 163). Examples of sentences where subjunctive/indicative variability is found in spoken Quebec French are given in (223) and (224), both taken from Poplack (1991, p. 236). In the sentences presented, the subjunctive and indicative occur under the same governor, *croire* “believe” and *admettons (que)* “let’s admit (that)” respectively.

(223)

- a. Mais quand tu es jeune, moi je crois que c' **est** une- une bonne chose.

“But when you're young, I think that it's (PRS IND) a good thing.”

- b. Je crois pas que ce **soit** la fin du monde.

“I don't think that it would be (PRS SUBJ) the end of the world.”

(224)

- a. Admettons mes deux petits **soient** détachés, ils peuvent me donner une amende pour deux fois vingt-huit piastres.

“Say both my kids are (PRS SUBJ) unbuckled, they can give me a ticket for two times twenty-eight bucks.”

- b. Admettons qu'elle **peut** pas ou que l'enfant **est** malade, on peut pas l'envoyer.

5 *A synchronic analysis of mood variability: the parameters that affect mood choice*

“Say that she can't (PRS IND) or that the kid is (PRS IND) sick, we can't send him .”

The three corpora on which the study is conducted represent three different temporal stages of spoken Quebec French. The first was collected in the 1940s and 1950s from elderly speakers born between 1846 and 1895 (Poplack & St-Amand, 2007). The second corpus was collected in the 1980s from speakers born between 1893 and 1964 (Poplack, 1989). The last corpus represents the informal talk of teenagers born between 1989 and 1991 and was collected in 2005 (Poplack & Bourdages, 2005).

The researchers draw on the apparent-time hypothesis described in section 5.1.4 of the present dissertation to analyze how the use of the subjunctive might have changed in over a century.

The statistical analyses assess the influence of different parameters, for instance the presence of elements that encode non-factual modality in the close syntactic context (adverbs like *peut-être* “maybe”, presence of the conditional or the imperative mood etc.). This parameter captures the hypothesis that the subjunctive is mainly an *irrealis* mood and should appear in non-factual contexts in opposition to the indicative. Also, purely morphosyntactic and lexical factors are considered, like the distance between the governor and the embedded verb or whether the embedded verb is irregular (*suppletive* in the authors' terms).

Over the years represented by the three corpora, the absolute occurrence of the subjunctive has actually increased. However, the

5 A synchronic analysis of mood variability: the parameters that affect mood choice

mood has become less productive. Its selection is increasingly constrained by the lexical and morphosyntactic context. As a matter of fact, the quantitative analyses show that the vast majority of subjunctives in subordinate clauses occur in combination with merely three governor verbs in spoken Québécois: *falloir* “be necessary”, *aimer* “like” and *vouloir* “want”. The subjunctive governor *falloir* accounts for 3/4 of all occurrences of subjunctive morphology and is thus labeled first-tier governor (Poplack et al. 2013, p. 165-167). Moreover, it was noted that the pool of embedded verbs that consistently take the subjunctive is not as varied either. Most subjunctives (65-67%) are found in only four verbs: *aller* “go”, *avoir*, “have”, *être* “be” and *faire* “do” (p. 167). These verbs are considered morphologically salient by the authors because their subjunctive form is highly irregular and is generated by unpredictable changes of the stem (eg. *aller* → *aille*.3/1.SG.PRS.SUBJ “go”, *faire* → *fasse*.3/1.SG.PRS.SUBJ “do”): “[...] morphological salience is a persistent predictor of subjunctive selection” (p. 167). On the other hand, “[...] regular embedded verbs [...] are far more likely to host other variants, presumably because they are less salient in such contexts [...]” (p. 172).

Based on these results, the researchers conclude that the subjunctive has been subject to fossilization and is progressively entrenched in fixed phrases. The lexical identity of the governor verb in combination with the lexical identity the embedded verb is the strongest predictor of the presence of the mood: “Operation of the lexical constraint does not

5 A synchronic analysis of mood variability: the parameters that affect mood choice

leave much room for purely productive use [...]” (p. 170). No semantic domain proposed in language grammars (e.g. *irrealis*) was found to predict the subjunctive in speech. Most importantly, the influence of the lexical identity of the governor and embedded verbs over the choice of the subjunctive increases significantly over the decades, peaking in the 21st century corpus where top-tier governors and top-tier embedded verbs combined came to account for 99% of all instances of subjunctive (p. 168).

Similar results are found for non-verbal governors. Only one non-verbal matrix, *pour que* “so that”, accounts for the majority of subjunctives in the three corpora, much in the same way as *falloir* among the verb governors. The subordinating locution is followed by *avant que* “before”, *mais que* “when”, and *jusqu’à tant que* “until” (pp. 176-177). Again, the lexical identity of a few elements proves to be the most important predictor of the subjunctives in speech.

Beyond the lexical influence of the governor element and the embedded verb, additional factors have been established which further constrain the environment of the subjunctive (see table 8 at p. 171 and table 9 at p. 173). The most significant are reported below.

Adjacency of governor to embedded verb

From the 21st century, the distance between the governor and the embedded verb starts to influence the selection of the subjunctive. Its probability decreases when the distance between the two elements increases, i.e. when additional linguistic material intervenes between

5 A synchronic analysis of mood variability: the parameters that affect mood choice

the two (see utterance (225) taken from Poplack et al. [2013, p. 1729] as an example).

(225) [...] il faut au moins tu te c –tu fais (IND) [for fasses (SUBJ)] pas mal à l'autre personne

(27)“[...] You have to at least – not hurt the other person.”
(20C.025.1114)

Presence of complementizer “que”

This factor begins to play a role from the 20th century onward, when the presence of the complementizer increases the probability that the subjunctive is selected.

Suppletive morphology in embedded verb

Irregular subjunctive forms derived through stem changes operated on the verb base correlate with the use of the subjunctive.

All these factors, along with the paramount lexical influence of the governor and the embedded verb point to the existence of prototypical subjunctive phrases or subjunctive collocations. They support the hypothesis concerning the mental representation of chunks presented in section 5.1.2.2: sequences of words that frequently occur together leave a trace in long-term memory. Their status as more entrenched units favors the preservation of obsolescing grammatical material. In those

5 A synchronic analysis of mood variability: the parameters that affect mood choice

structures, the subjunctive resists morphological levelling to the indicative.

Indeed, when the unit of the subjunctive phrase is broken, for instance, when additional linguistic items occur between the main and the embedded clause, the probability of selecting the subjunctive is reduced (see (225)).

The subjunctive phrases are made of four or five lexical items at least (see again experimental evidence from Arnon & Snider (2010) large phrases made of four words that are highly frequent in speech may leave a trace in memory). A prototypical French subjunctive construction looks like the following:

				<i>fasse</i>					
				<i>ait</i>		“it	is	necessary	that
(il)	<i>faut</i>			<i>soit</i>		do/have/be/go/can.3/1.SG.PRS.SUBJ”			
	<i>veut</i>	<i>que</i>	[SJCT]	<i>aille</i>		“he wants that [SJCT] do/have/be/go/can.3/1.SG.PRS.			
				<i>auisse</i>					

The researchers observe that the embedded verbs that most consistently select the subjunctive in speech are all high-frequency verbs in addition to being irregular (Poplack et al., 2013, p. 167). This evidence supports the hypothesis relating high token frequency to cognitive entrenchment according to Bybee & Thompson (1997, p. 381): frequent forms resist levelling to the more productive

5 A synchronic analysis of mood variability: the parameters that affect mood choice

morphological variant because they are entrenched in memory (see again section 5.1.2.2).

In the corpus studies on Quebec French, the embedded subjunctive does not consistently encode the semantic functions usually associated with it, i.e. subjective, non-declarative or irrealis scenarios. It neither occurs reliably with governor verbs of semantic classes that are supposed to echo the semantics of the mood, such as directive or optative predicates, as Farkas (1992b, p. 72-74) had suggested (see again section 3.1), nor is it triggered by contextual elements (i.e. adverbs, locutions etc.) that express non-factual modality. It becomes a subordination marker devoid of any distinctive meaning. This evolution into a syntactic marker takes place through desemanticization and generalization of its functions: as the subjunctive is found in increasingly diverse clauses, a common denominator can no longer be inferred. These processes must have already started even before the period of Classical Latin where the embedded subjunctive no longer functioned as a modal marker with a specific meaning according to the analysis conducted by Pinkster (2015, p. 619) and presented in section 4.3.

Probably, the fact that the embedded subjunctive does not convey any additional meaning to the sentence (Blücher, 2003; Porto Dapena, 1991) and that speakers cannot associate the subjunctive with stable semantics makes other moods like the indicative equally possible in subjunctive clauses. For this reason, the subjunctive suffers from competition with the indicative, which, being by far the most

5 A synchronic analysis of mood variability: the parameters that affect mood choice

productive and most frequent morphological system, tends to replace it with the exception of the lexical items and phrases that are rooted in memory.

In Poplack et al. (2018), the authors extend the corpus study of the subjunctive previously conducted on Quebec French to the other Romance languages, Spanish, Portuguese and Italian. The objective of the analysis is to establish the degree to which the subjunctive is grammaticalized in each language by looking at the processes that accompany grammaticalization: desemanticization and obligatorification or conventionalization (processes presented in section 5.1.1). Grammaticalization advances when the subjunctive is increasingly associated with elements of the close syntactic context, a phenomenon variously called obligatorification or entrenchment and is no longer a semantically motivated choice (desemanticization).

To test whether the subjunctive occurs as an indicator of a certain modal force, similar to the previous studies, the matrix clause of the sentences displaying mood variation was analyzed for the presence of linguistic elements that can be associated with non-assertive, conditional or interrogative illocutionary force on the one hand (e.g. negation particles like “not”, “never”, conjunctions like “if”, interrogative sentences) and for elements conveying non-factual modality on the other (e.g. epistemic adverbs like “maybe”, “potentially”, evaluative adverbs “hopefully”, modal elements “it may be”, or the presence of certain tenses and moods such as future,

5 A synchronic analysis of mood variability: the parameters that affect mood choice

conditional). It was then established whether the selection of the subjunctive correlated with those elements as opposed to the indicative.

The subjunctive was shown to correlate with these indicators of non-factual, non-assertive, conditional modality and with polarity only in Spanish (pp. 14-15). In the other languages no particular correlation can be established. The authors suggest that in French and Italian negative, conditional or interrogative contexts actually seem to inhibit it (p. 15).

Further, it was explored whether the semantic class of the governor verbs bears a significant relation to the subjunctive. First of all, no governor verb was found to categorically select the subjunctive in speech in either French, Portuguese or Italian, not even the prototypical volitive verbs, that were taken by many researchers to express the core subjunctive meaning (see again section 3.1). In Spanish, however, the volitive class bears a more systematic relation to the subjunctive which is selected categorically or nearly categorically with verbs like *querer* “want”, *gustar* “please”, *dejar* “permit” or *hacer que* “make-causative”. Interestingly, verbs belonging to the same class, synonymous or nearly synonymous verbs bear very different correlations with the subjunctive, in Spanish as well, with the exception of the afore-mentioned volitive class. This is especially the case for epistemic (e.g. “suppose”, “imagine”, “think” “believe”) and factive/emotive governors (e.g. “be happy”, “be certain”) (pp. 16 -18).

Based on the data presented, the authors conclude that the subjunctive is highly variable in all circumstances evaluated and in all languages. The semantic functions that are usually associated with the

5 A synchronic analysis of mood variability: the parameters that affect mood choice

subjunctive are the encoding of non-assertive (Bolinger, 1974, p. 465; Thompson, 1998; Chafe, 1995; Givón, 1978) and non-factual and subjective modality (Givón 1994, p. 271). None of them can account for the subjunctive/indicative alternation in the languages examined with the exception of Spanish, where linguistic elements relating to negative polarity, non-factuality, conditional and interrogative modality seem to favor the subjunctive.

The authors note that in French, Italian and Portuguese the subjunctive mood appears emancipated from the semantic functions that it is assumed to convey. Spanish unlike the other languages is more conservative in this respect. Semantics still plays a role in the subjunctive / indicative distribution although variability occurs (Poplack et al., 2018, 18).

Having established the degree of desemanticization in the Romance languages, the authors investigate to what extent the subjunctive is routinized in certain phrases. In particular, they examine the lexical identity of the governor verbs which was found to play a major role in the subjunctive/indicative variability in the previous studies on Quebec French. In all the languages considered, out of all the governor verbs that select the subjunctive at least once, only a very small number are frequent. In French, these are *falloir* “be necessary” and *vouloir* “want”, in Portuguese *querer* “want”, in Spanish *decir* “say”, *creer* “believe”, *querer* “want”, in Italian *pensare* “think”, *credere* “believe”, *sembrare* “seem”. Also, a very small number of matrix verbs, two or three, strongly correlates with the subjunctive (among them, the equivalent of

5 A synchronic analysis of mood variability: the parameters that affect mood choice

verb “want” figures as a strong predictor of subjunctive morphology in all languages), other verbs exhibit a high degree of mood variation. A third group of governors consistently favors the indicative instead. They are not the same across languages. The associations are highly idiosyncratic and, as was discussed before, do not generalize over semantic classes. In French, the lexical effect of the governor is particularly striking: one verb (*falloir* “be necessary”) accounts for nearly 80% of all the subjunctives occurring in the corpus. The majority of the remaining subjunctives occurs with *vouloir* “want” (p. 21).

Additionally, lexical constraints coming from the embedded verbs were established. Only a handful of predicates were found to occur with more than half of the subjunctives in Italian and French. They are all high-frequency irregular verbs (mainly auxiliaries and modals). They account for approximately 35% of the total subjunctives in Portuguese and Spanish (p. 22).

Unfortunately, no statistics are presented concerning the correlation of the embedded verbs with the subjunctive as opposed to the indicative or the conditional in the variation loci considered. Thus, for instance, it is not possible to determine whether verb auxiliaries “be” or “have” in the various languages indeed prefer the subjunctive or equally select the indicative. It remains to be investigated how and to what extent the lexical identity of the embedded verb constrains the morphological variation, and more specifically, whether frequency and morphological irregularity plays a role in the mood distribution. Finally, the distribution of subjunctives across verbs becomes a more meaningful

5 A synchronic analysis of mood variability: the parameters that affect mood choice

statistic if it is compared to the distribution of indicatives or other moods across verbs. In this way, by comparison to a baseline, it can be established whether the subjunctive differs from other moods in that only few highly frequent verbs occur with it in the percentages presented. Given the Zipfian distribution (Zipf, 1999) of words according to their frequency (the most frequent word occurs roughly twice as often as the second most frequent word), a comparison remains necessary.

Based on the analysis presented, the authors observe that the subjunctive is subject to lexical effects: the lexical identity of the governor verb greatly predicts the subjunctive along with the lexical identity of the embedded verb. These effects are particularly strong in French, but also in Italian. The subjunctive thus appears to be routinized or entrenched in specific phrases. The degree of entrenchment varies across languages: Spanish again emerges as slightly less affected than the other idioms. The routinization of the subjunctive in a few formulaic schemas along with its desemanticization and its overall low occurrence rate leads to the conclusion that the mood is no longer truly productive in French, Italian and Portuguese. It can be positioned at the late stages of grammaticalization and may eventually completely disappear (see Figure 5.1 depicting the various stages of grammaticalization ending with complete loss of the grammatical material). Spanish lags slightly behind on the grammaticalization cline.

As the studies reviewed show, corpus research provides valuable insights into mood variation and enables the researcher to test very fine-

5 A synchronic analysis of mood variability: the parameters that affect mood choice

grained hypotheses concerning morphological variation and productivity. However, one must acknowledge the limitations of the speech corpora available in the Romance languages for the analysis of the subjunctive. For a quantitative analysis whose results need to be predictive of future linguistic behavior, the corpora available are small and collect very few occurrences of the subjunctive overall. Indeed, the analysts in many cases dealt with a very restricted number of subjunctives and indicatives in a particular variation locus. For instance, under certain governors, the absolute numbers of the subjunctives and indicative taken together were below 10 or 20. Such numbers hardly warrant any generalizations which would be prone to a high degree of error and can simply reflect chance. Moreover, such corpora cannot host an analysis which examines the effect of multiple predictors at once. Considering for instance the combined effect of a certain governor and a certain embedded verb on mood variation further reduces the instances where the subjunctive is found under these particular conditions. In most cases there would simply be no observations.

Yet, studying the extent to which the subjunctive is routinized in certain phrases requires mapping additional syntactic factors which may be equally important in determining mood variation. Some of those are the verb person, verb tense, the presence vs. absence of the complementizer etc.

The corpus study that I present in section 5.3 suffers from the same limitations. In order to overcome the limits due to a relatively small

5 A synchronic analysis of mood variability: the parameters that affect mood choice

collection of data for the phenomena investigated, corpus research benefits from being complemented with experimental designs. The latter offer the advantage of directly testing the variable(s) targeted by the researcher while keeping other variables constant and ideally gather sufficient observations in order to produce statistically valid generalizations.

Quantitative studies within the variationist framework have examined the evolution of the subjunctive in Spanish in situations of language contact (Silva-Corvalán, 1994a, 1994b; Montrul, 2007, 2009). Language contact is considered particularly interesting because it is found to accelerate processes of grammaticalization and language change due to the rapid restructuring of the social tissue of the linguistic communities and networks (Thomason, 2001; Heine & Kuteva, 2005; Matras, 2011). Changes relating to attrition and loss usually take place in a shorter amount of time when the new generation of speakers is no longer in contact with their native language in a variety of communication settings which are taken over by other languages. Silva-Corvalán (1994b) analyzes the use of the subjunctive in the conversational talk of three generations of Mexican-American bilinguals coming from immigrant families and living in Los Angeles. She draws on the apparent-time hypothesis and compares the language behavior of the younger generations to that of the older one to gather patterns of morphological simplification. Of the eighteen syntactic contexts in which the variable use of the subjunctive was analyzed, only the causative construction prompted the subjunctive categorically (see

5 A synchronic analysis of mood variability: the parameters that affect mood choice

sentence (226)) (1994b, p. 264). This result parallels the analyses by Poplack et al. (2018) according to which the subjunctive is always selected under causative governor *hacer que* “make so that” in Spanish.

(226) Hice que todo fuera (IMPF SUBJ) más difícil.

“I made everything more difficult.” (abrinegocio.com, Corpus del Español: Web/Dialects)

Overall, the subjunctive rates are found to significantly decrease from the first to the third generation of speakers. The syntactic contexts where the subjunctive remains most stable are volitional, purpose clauses, and concessive clauses. The author does not give any data on the number of governors or lexical elements that represent each syntactic/semantic context. Thus, we cannot make any observation on the lexicalization rate of the subjunctive over the three generations. The study, however, draws a comparison between the contexts where the subjunctive is obligatory (i.e. lexically selected by the governor) and those where it is optional, reporting that the rate of decline of the subjunctive is overall greater in the latter. Such results parallel others (Klein-Andreu, 1980; Ocampo, 1990; Montrul, 2009) and according to the researcher indicate that the subjunctive/indicative opposition progressively loses its contrastiveness. The subjunctive is still selected lexically by a number of governors but no longer conveys modal distinctions, a conclusion shared with the studies presented above (Poplack, 1990; Poplack et al., 2013; Poplack et al., 2018).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

A number of studies have investigated whether mood variability is conditioned by verb regularity in the speech of Spanish native speakers and L2 speakers (Ross Veidmark & Umaña Aguiar, 1991; Lubbers, 1998; Geeslin & Gudmestad, 2008; Gudmestad, 2010, 2012a, 2012b). The corpus analyses by Poplack et al. (2013) and Poplack et al. (2018) previously discussed had shown a special relation between the subjunctive and irregular verbs. However, the works that explicitly target form regularity as a factor report inconsistent findings largely dependent on the experimental design and on the operationalization of the regularity/irregularity distinction.

5.3 Effects of lexicalization on mood variability in Italian: a corpus study

In the corpus studies reviewed above a number of important findings have been discussed. Mood variation in the Romance languages must be understood as a synchronic manifestation of processes connected to language change and grammaticalization: specifically, obligatorification and desemanticization. The occurrence of the subjunctive in embedded clauses cannot be explained by relying solely on semantic distinctions that the speakers intend to encode, not even in Spanish where semantic contrasts connected to modality still play a role. Instead, a number of lexical and morpho-syntactic factors influence the selection of mood in speech. First and foremost, the lexical identity of the governor to a large extent predicts the occurrence rate of the subjunctive and the indicative. The embedded verb is also a possible

5 A synchronic analysis of mood variability: the parameters that affect mood choice

predictor of subjunctive morphology. In the previous studies, a large portion of the subjunctives (60%-70%) was found to occur with only a few verbs in French and in Italian. They are all high-frequency irregular verbs. Thus, frequency and regularity emerge as further possible variables influencing mood variability, although the studies explicitly targeting regularity in Spanish have yielded contradicting results. Frequency, in particular, is assumed to act as a conserving force against analogical leveling in language change: high-frequency lexical tokens usually resist the morphological shift to the new more productive morphological variant (Bybee & Slobin, 1982; Lieberman et al., 2007).

Additional predictors were found to influence the selection of the subjunctive in French: presence of the complementizer *que* “that” and the distance between the governor and the embedded verbs among others. All these factors taken together lead to the conclusion that the subjunctive is in fact not entirely productive (the degree of productivity varies according to the language in question). It is rather entrenched in specific subjunctive schemas or phrases (Bybee & Thompson, 1997, p. 383). Elements of the close syntactic contexts prompt its selection. The correlation between the subjunctive and other linguistic elements may appear idiosyncratic. With regard to the mood variation in Italian, in a previous experiment (Pietropaolo, 2014), I showed how verb person can also be a strong predictor of the mood selected. In particular, the subjunctive appears systematically disfavored when the embedded verb is conjugated in the second person singular.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

In the corpus study presented in the following sections, I analyze the role of the governor verb on the selection of subjunctive morphology in embedded clauses and establish whether the lexical semantics of the governor can account for the mood distribution as suggested by semantic-based accounts (Farkas, 1992b; Giorgi & Pianesi, 1997 among others).

Additionally, I explore the distributional properties of the subjunctive among verb types and compare them to those of the infinitive. The infinitive mood was chosen as a baseline for the comparison instead of the indicative or other mood forms because it has a clear productive pattern (infinitive suffix *-are*), it has been researched in the literature on morphological learning and productivity (Albright, 2002a; Veríssimo & Clahsen, 2009, 2014) and serves as a point of contact with previous studies. The distributional properties of the two moods will serve as the basis for the study on the factors influencing morphological productivity presented in chapter 6.

5.3.1 Materials for the corpus analysis: The C-ORAL-ROM corpus

The study uses the Italian language sample provided by the C-ORAL-ROM corpus (Cresti et al., 2002; Cresti & Moneglia, 2005; Moneglia, 2003). The C-ORAL-ROM gathers four corpora of spoken Romance Languages, Italian, French, Spanish and Portuguese. Roughly 300,000 lexical tokens are collected for each language. The Italian collection is taken from the LABLITA corpus (Cresti, 2000; Moneglia, 2005) and

5 A synchronic analysis of mood variability: the parameters that affect mood choice

compiles spoken texts from Western Tuscany. It represents a valuable resource for studies on morphology because it is transcribed and automatically annotated at the word level, at the lemma level and at the part-of-speech level (PoS). It gives very detailed morphological information about each word. For instance, in the case of verbs, the tags report whether the verb form in question is a finite or non-finite form and the verb mood (whether it is an indicative, a subjunctive, an imperative etc.). It additionally gives information about the verb tense and person (see Cresti & Moneglia [2005 pp. 88-91] for a comprehensive presentation of the tagset employed in the annotation of verbs in the Italian C-ORAL-ROM).

The Italian corpus contains adult speech in dialogue and monologue format produced in both informal and formal settings. The language sample gathered in formal settings is comprised of political speeches and debates, news and legal talks among others. The informal talk contains private dialogues and conversations with friends and family. The inclusion of formal speech in the corpus has non-trivial consequences for the variable under investigation. It was discussed in chapter 2 how mood variation is inevitably affected by the social context in which communication takes place. Monologues and partially scripted talk occurring in very formal contexts and reflecting high, prestigious registers more rigidly conform to the prescriptive canons of the standard variety and avoid stigmatized linguistic practices. Accordingly, the subjunctive, as it is still considered to be the standard and correct variant in the clauses analyzed, will be preferred even if

5 A synchronic analysis of mood variability: the parameters that affect mood choice

instances of mood variation are found in these registers (see again chapter 2 for a discussion of morphological errors in controlled speech registers). Speakers in informal colloquial settings, on the other hand, are more spontaneous and least affected by concerns of prestige and correctness. Informal speech thus represents the ideal candidate for the study of mood variation. Unfortunately, the metadata concerning the speech register (formal vs. informal) and the speech type (scripted monologue vs. talk) was not available in the corpus. For this reason, it was not possible to control for socio-pragmatic dimensions of variation. These represent serious confounds on a study testing the role of frequency on the occurrence of the subjunctive: very infrequent lexical items related to technical jargon or a very formal register may appear to prefer the subjunctive over non-standard morphological forms. For this reasons, in section 5.3.3.2, I limit myself to a descriptive and exploratory data analysis on the distribution of subjunctive forms among verb types and tokens. The influence of the governor on the probability of selecting the subjunctive will be explored using inferential statistics where possible confounds coming from the lexical identity of the embedded verbs and their frequency will be controlled for.

5.3.2 Method

The first challenge to the study of mood variation is to identify all the contexts where the subjunctive freely alternates with other moods. I adopted the empirical approach described in Poplack (1990) and

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Poplack et al. (2013). First of all, the subjunctives were extracted from the corpus. This was possible because the transcribed corpus is annotated at the PoS level as previously discussed. The *Contextes* text search engine provided by the C-ORAL-ROM was used to find all the subjunctives tagged in the corpus. During the extraction phase it was noticed that the morphological annotation provided by the C-ORAL-ROM corpus was in fact missing from certain lexical items, mainly irregular verb forms, in the POS-annotated transcriptions. For this reason, in a second step, regular expressions (regex) were used to extract the subjunctive forms.

In Italian, the subjunctive and the indicative forms happen to be syncretic and thus cannot be distinguished in two contexts: 2nd person singular for the verbs of the 1st conjugation (e.g. tu parli ‘you speak’), and 1st person plural across all three conjugations (e.g. noi parl-iamo [1st conjugation] ‘we speak’, fing-iamo [2nd conjugation] ‘we pretend’, part-iamo [3rd conjugation] ‘we leave’). These forms were excluded from the analyses.

The syntactic context in which each subjunctive occurred was analyzed and the subjunctives were divided into two groups: those appearing in the main clause (unembedded subjunctives) and those appearing in the subordinate clause (embedded subjunctives). The latter group was further studied to identify the governors that license the subjunctive clause. Overall, 75 verbal and non-verbal governors were found to select the subjunctive at least once. They were considered the environment where the subjunctive could at least in theory be

5 A synchronic analysis of mood variability: the parameters that affect mood choice

substituted by other mood forms without generating a functional distinction, i.e. the locus of mood variation. Thus, in a second step, all the verbs controlled by the governor set were collected and their mood form was annotated.

Of the original 75 governors, 15 were selected as the locus of variation for the current analysis, the governors which were found to have the highest rates of occurrences in the corpus.¹⁰ This yielded 350 subjunctive/indicative (or conditional) observations.

In order to investigate to what extent each governor influences the probability of selecting the subjunctive, the lexical identity of the governors was entered as a categorical predictor variable into a logistic regression (GLMMs) (Bates et al., 2014). The mood form was considered the dependent variable in the statistical analysis. It was coded as a binary variable: it could take two values, presence of subjunctive = 1, absence of subjunctive = 0.

The probability of occurrence of a subjunctive under each governor was contrasted with a probability distribution in which the two outcomes, presence vs. absence of a subjunctive, are equally likely (probability of 0.5 for each of the two outcomes). This is a maximum-entropy distribution¹¹ and represents an ideal state of inherent variability

¹⁰ The other governors were excluded from the analysis because they occurred overall only once or twice in the entire corpus. Thus it was impossible to calculate reliable estimates of their correlation with either mood form.

¹¹ The maximum-entropy distribution can be used to describe the tosses of a coin. If a coin is fair (unbiased), the two sides of a coin (heads or tails) have equal probability of occurrence. Thus there is the highest uncertainty concerning the outcome of the toss.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

between the subjunctive and the indicative: one of the two mood forms does not occur more often than the other. The maximum-entropy distribution is encoded in the intercept of the model to which the mood distribution under each governor is compared.

The model has as control variables the lexical identity of the embedded verb and its token frequency per million. Token frequency was log-transformed, centered, and standardized before entering it as a control variable in the model.

The model shows which governor significantly increases the probability of selecting a subjunctive compared to a state in which the two moods are equally likely (Appendix A reports the coefficients of the model and the model formula). Based on the data, it will be evaluated whether governors that can be grouped together in the same semantic class according to their lexical meaning exhibit comparable subjunctive rates, a hypothesis suggested by semantic-based theories (Farkas, 1992b; Giorgi & Pianesi, 1997), that see the prime factor driving the occurrence of the subjunctive in the meaning of the governor (see again chapter 3).

The statistical analysis was conducted with R Studio (Version 0.99.896, 2009-2016 RStudio, Inc.) (Allaire, 2012) using the `glmer` function of the `lme4` package (Bates et al., 2007; Bates et al., 2012).

Finally, a descriptive analysis of the distribution of subjunctive forms among verb types and in particular among high-frequency verbs is presented. The proportion of subjunctives is compared to the

proportion of infinitive forms. All the subjunctives (overall 1148 tokens) and the infinitives (overall 8925 tokens) extracted from the corpus were included in this analysis.

5.3.3 Results

5.3.3.1 The influence of the governor verb on the subjunctive/indicative distribution in speech.

Figure 5.2 shows the distribution of subjunctives and indicatives by verb governor. The plot reveals the pervasiveness of mood variation in speech: no single verbal or non-verbal governor selects the subjunctive categorically (a similar result is found in Poplack et al. [2018]). The rate of occurrence of the subjunctive varies according to the governor that appears in the matrix clause. However, only two governors significantly increase the probability of selecting a subjunctive according to the estimates of the model (see Figure 5.3). These are *bisognare* ‘to necessitate’ and *volere* ‘to want’. They are the only two governors that consistently select the subjunctive in speech. Incidentally, the same two governors (il) *faut* ‘to necessitate’ and *vouloir* ‘to want’ were found to be the only ones to reliably predict subjunctive morphology in French (Poplack et al., 2013).

One governor *non essere che* ‘to be not that’ correlates with the indicative. It encodes the meaning of negation (retracting an assertion), which according to semantic accounts (Quer, 1997, p. 171) is one of the primary functions of the subjunctive in the Romance languages (see again the notion of polarity subjunctive discussed in chapter 3).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

All the other governors show inherent variability. The subjunctive/indicative distribution under these governors does not significantly differ from the condition in which the two moods are equally likely (see Figure 5.3). In fact, not one of these governors significantly increases the probability of selecting a subjunctive. They are all expected to occur with subjunctive morphology according to semantic accounts and prescriptive grammars because they encode modal values associated with the subjunctive mood (Farkas, 1992b). For instance, verb *sperare* “hope” conveys an optative meaning; governors *sembrare* “seem”, *parere* “seem”, *credere* “believe”, *pensare* “think” are prototypical expressions of epistemic or evidential meanings; non-verbal governors *se* “if”, *nel caso (che)* “in the event (that) / if”, and verb governor *potere darsi* “may be” refer to irrealis or contrary-to-fact situations that are not actualized in the real world, the quintessential domain of the subjunctive according to semantic accounts (see Farkas, 1992b; Givón, 1994; Giorgi & Pianesi, 1997 among others).

5 A synchronic analysis of mood variability: the parameters that affect mood choice

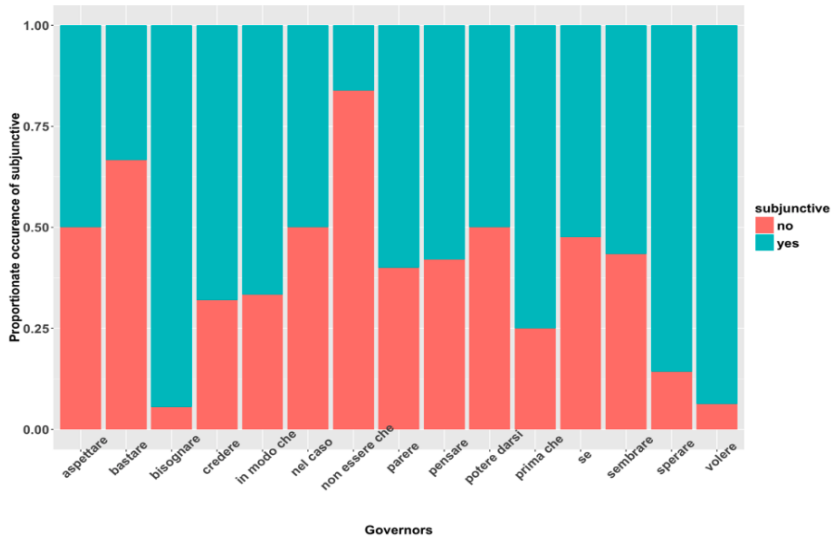


Figure 5.2: Distribution of subjunctive forms by governor

5 A synchronic analysis of mood variability: the parameters that affect mood choice

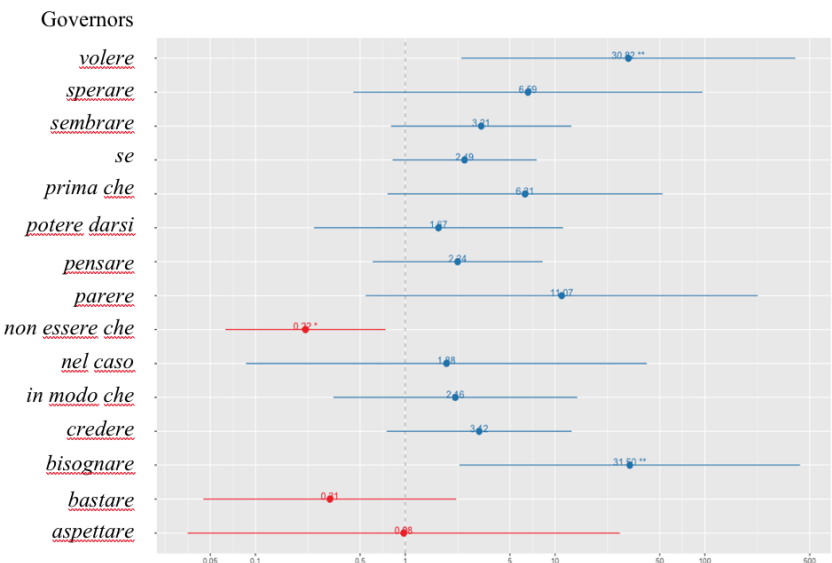


Figure 5.3: Effect of governor on the probability of occurrence of a subjunctive. Estimates of the logistic regression with 95% confidence intervals.

Crucially, quasi-synonymous governors or governors that belong to the same semantic class exhibit different rates of subjunctive occurrences. Thus, *credere* “believe” occurs more often with a subjunctive than quasi-synonymous verb *pensare* “think”. Governors *bisognare* “be necessary” and *bastare* “be sufficient” are arguably semantically similar in that they express a necessary condition and for this reason they select the subjunctive according to DellaValle & Patota (2011, pp. 108-109). Yet, the former favors the subjunctive while the latter favors the indicative in speech.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Also, semantic accounts assume that the intentional subjunctive, which appears under volitive and directive predicates, encodes the core subjunctive meaning and for this reason should resist mood variation both within one language and crosslinguistically (Quer, 1997, 2009; Kempchinsky, 2009). However, only verb governor *volere* “to want” appears to correlate with the subjunctive. No other governor encoding directive or volitive modality appears with the subjunctive. The semantic domain pertaining to deontic/volitive modality is represented by a single matrix.

According to the evidence gathered, the opposition between the subjunctive and the indicative based on modality proposed by the literature reviewed in chapter 3 reflects an idealized state of the grammar rather than the speakers’ internal linguistic knowledge as it emerges from the empirical observation of actual speech.

The idiosyncratic correlation between governors and the subjunctive strengthens the hypothesis that the mood distribution in subordinate sentences does not respond to a semantic principle. The fact that only two governors consistently select the subjunctive supports the hypothesis that the mood has lexicalized in a few prototypical phrases which resist morphological variation (Bybee et al., 1994; Poplack et al., 2013).

5.3.3.2 The mood distribution among high-frequency verbs: a comparison between the subjunctive and the infinitive mood

Table 5.2 shows a comparison between the distribution of subjunctive and infinitive morphology among the verbs with the highest frequency and with the highest number of subjunctive tokens in running speech in the C-ORAL-ROM. The infinitive, as a productive mood, serves as a baseline.

It is possible to note that five high-frequency verbs *essere* “to be”, *avere* “to have”, *potere* “can”, *fare* “to do”, *venire* “come” make up only 21.2% of the infinitive occurrences whereas they amount to 60.8% of subjunctive occurrences, a striking difference. The latter statistic is in line with the data on the Italian subjunctive presented by Poplack et al. (2018, p. 22). The majority of subjunctive forms, unlike the infinitive, are selected by a few high-frequency forms. Especially the verb *essere* “to be” exhibits a disproportionate rate of subjunctive tokens with respect to the other verbs or with respect to its infinitive rates.

Finally, Table 5.3 and Table 5.4 display the overall number of verb types (i.e. the unique verbs), the overall number of tokens and the type-token ratio of the subjunctives and the infinitives in the C-ORAL-ROM.

Type-token ratio is one of the simplest measures of lexical richness, or diversity in vocabulary. In our case, it shows how varied the verbs that occur with the subjunctive and the infinitive are. It is obtained by dividing the total number of unique words (types) by the total number

5 A synchronic analysis of mood variability: the parameters that affect mood choice

of word occurrences (tokens) (Johnson, 1944). Overall, the type-token ratios of the subjunctive and infinitive do not differ much. The subjunctive actually scores a little higher on the lexical diversity measure, 0.13 vs. 0.11 for the infinitive. This is because it is very infrequent overall, especially among low-frequency verbs of the 1st conjugation. Indeed, both the number of verb types and the overall tokens that occur with the subjunctive are extremely low with respect to the infinitive.

Most interestingly, the subjunctive and the infinitive exhibit different distribution rates among the conjugation classes. The infinitive mood is most frequent among verbs of the 1st conjugation (4492 tokens and 668 types). These rates match the size of the conjugation classes in Italian: the 1st conjugation is by far the largest and the only productive class in Italian. It includes the great majority of verbs (approximately 75%) and it is used to derive new verb forms (e.g. *messaggiare* “to message” from noun *messaggio* “message”). On the other hand, the subjunctive occurs most often among verbs of the 2nd conjugation (710 tokens), a non-productive class that overall groups together far fewer verbs. Most high-frequency irregular verbs belong to this class.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Table 5.2

Distribution of subjunctive and infinitive morphology among high-frequency verbs in C-ORAL-ROM

Verbs		Subjunctive	Infinitive
<i>essere</i>	‘be’	34.8%	6.8%
<i>avere</i>	‘have’	10.2%	2.4%
<i>potere</i>	‘can’	7.1%	0.3%
<i>fare</i>	‘do’	5.1%	11.0%
<i>venire</i>	‘come’	3.6%	0.7%
<i>sentire</i>	‘listen’	3.0%	0.7%
<i>dire</i>	‘say’	3.0%	7.7%

Table 5.3

Type/token distribution of infinitives in C-ORAL-ROM

	Types	Tokens	Type-token ratio
overall	1009	8925	0.11
1 st conjugation	668	4492	0.14
2 nd conjugation	191	2592	0.07
3 rd conjugation	129	1752	0.07

5 A synchronic analysis of mood variability: the parameters that affect mood choice

Table 5.4

Type/token distribution of subjunctives in C-ORAL-ROM

	Types	Tokens	Type-token ratio
overall	156	1148	0.13
1 st conjugation	83	276	0.30
2 nd conjugation	37	710	0.05
3 rd conjugation	36	162	0.22

5.3.4 Conclusion

By means of a corpus analysis it was shown that the subjunctive and the indicative alternate freely in speech in subordinate clauses governed by the same lexical element. The rates of subjunctive forms under each governor vary in a way that appears arbitrary. No particular patterns emerge that can be related to the lexical semantics of the governor. Indeed, synonymous verbs or verbs belonging to the same semantic class exhibit different correlations with the subjunctive or the indicative. Crucially, these correlations do not significantly differ from a condition in which the two moods are equally likely. The only two governors which significantly correlate with the occurrence of the subjunctive are *bisognare* “to necessitate” and *volere* “to want”. This result is in line with the findings reported in Poplack et al. (2018) who use only descriptive statistics for their corpus analysis.¹² The evidence

¹² The present statistical analysis was conducted a year before the paper authored by Poplack et al. (2018) was published.

5 A synchronic analysis of mood variability: the parameters that affect mood choice

suggests that the subjunctive must have been routinized in these particular phrases (see again the process of obliigatorication discussed in section 5.1.1 and the notion of subjunctive schemas presented in Bybee et al. [1994]). The mood distribution in the corpus contradicts the accounts according to which a semantic variable governs the subjunctive/indicative distribution in subordinate clauses.

Finally, it was shown that the subjunctive, unlike the infinitive mood, is very infrequent in discourse and occurs for the most part with very few verbs (only five lexical items), a fact that suggests low productivity. Indeed, most subjunctive forms encountered in the corpus belong to the 2nd conjugation, a class that groups together very few high-frequency verbs, for the most part irregular. This conjugation is considered to be non-productive in Italian. Thus, subjunctive morphemes of the 2nd conjugations appear to be the most frequent in speech. This is not the case for the infinitive, for which the morphological pattern of the 1st conjugation is the most common in speech.

This peculiar distribution of the subjunctive among verb classes will be used to explore the relationship between frequency and the formation of productive rules in chapter 6. Additionally, the study will investigate the productivity rates of the different subjunctive patterns.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

The present chapter investigates the productivity of regular and irregular subjunctive patterns and examines a potential relation between the frequency of occurrence of each pattern and its productivity with novel words. In particular, the following research questions will be addressed:

- a) How do speakers generalize morphological patterns to new words?
- b) Is the generalization of both regular and irregular morphological patterns influenced by similarity to existing words?
- c) Does frequency of occurrence of the morphological patterns influence their productivity?

In the past decades, three main theoretical proposals have been put forward: single-route rule-based models (Marcus et al., 1992), dual-mechanism models (Pinker, 1991) and analogy-based models (Rumelhart & McClelland, 1986; Bybee, 1995). The models disagree on the mechanisms that drive generalization of morphological patterns and their productivity.

For instance, the dual-mechanism model makes “an explicit distinction between context-free and similarity-based generalizations” (Veríssimo & Clahsen, 2014, p. 61). It posits that regular morphological processes (e.g. love > loved) are applied to new words by means of a context-independent rule. This mechanism allows unbounded generalizations to

6 The learning and generalization of variable morphological rules: the case of the subjunctive

words that have never been heard before and that do not resemble anything that the speaker has experienced so far. For example, if a hypothetical new verb *to soof* were to be adopted in the English language, native speakers would easily be able to generate its morphologically complex variants (e.g. *soofed*, *soofing*). On the other hand, novel use of irregular morphology (e.g. *bite* > *bit*) is rare and is prompted by phonological similarity to attested examples (Prasada & Pinker, 1993).

Contrary to this hypothesis, according to analogy-based models, morphological generalization is induced first and foremost by similarity to known words (semantic and phonetic). Rather than applying one rule to all members of a category by default, speakers produce probabilistic generalizations. The probability that a morphological pattern will be selected depends on its frequency in the linguistic input that the speaker receives (Rumelhart & McClelland, 1986; McClelland & Patterson, 2002; Bybee, 1995).

The hypotheses of the different models are tested using an oral production task in which adult native speakers of Italian are asked to derive subjunctives and infinitives of new words. The infinitive, having a clear productive pattern (see again table Table 5.2) serves to establish a baseline in order to interpret the data and the patterns that emerge from the analysis.

So far, no study exploring the above-mentioned hypotheses has been conducted on the subjunctive. Also, studies investigating the use of the subjunctive have always attempted to give an overall measure of its

6 The learning and generalization of variable morphological rules: the case of the subjunctive

productivity. To the best of my knowledge, no other work has analyzed the productivity rates of each subjunctive pattern.

The subjunctive mood is especially well suited to test the hypotheses on the influence of frequency on morphological productivity. Indeed, as it will be shown in section 6.5, in comparison to other moods like the infinitive, the subjunctive exhibits a larger group of verbs that take the irregular pattern. Also, irregular subjunctive forms are especially frequent in speech (see again Table 5.2). For this reason, the irregular subjunctive pattern has a higher chance of attracting new words. Having a more balanced distribution of regulars and irregulars improves the possibility of detecting graded effects of frequency and similarity.

6.1 Models of morphological learning and productivity

Long-standing research questions in psycholinguistics are how morphological rules are learned and how complex forms are produced and processed. Rule-based models and analogy-based models provide different explanations which range from obligatory decomposition of words to storage of all complex words. The debate also concentrates on possible differences between regular inflectional processes (e.g. English past tense *help* > *helped*) and irregular processes (e.g. *do* > *did*, *go* > *went*). Three main theories are introduced and discussed:

- 1) single-route rule-based models
- 2) dual-mechanism models
- 3) single-route analogy-based models

6 The learning and generalization of variable morphological rules: the case of the subjunctive

According to single-route rule-based models all complex forms, regular and irregular, are generated by a rule. Rules are formally distinct from words. They are understood as operations conducted over words and are learned as “paired associations” between the base and the derived form (Marcus et al., 1992; Marcus, 2001). Whether a rule will be used with novel words or not depends on whether it is productive in the language considered. The account does not foresee a role played by frequency of use in rule formation. Baayen (1992, p. 138) notes that, although Anshen & Aronoff (1988) acknowledge that token frequencies might be relevant, word formation rules as presented in Aronoff (1976) cannot incorporate frequency effects in their formal representation.

Another important position has been developed by the proponents of the dual-mechanism model (Pinker, 1991, 1999; Pinker & Prince, 1988, 1994; Prasada & Pinker, 1993; Clahsen, 1999). They make a fundamental distinction in the way regular and irregular morphological processes are learned and extended to new forms. Focusing on the English past tense, they observe that the regular past tense inflection with segmentable suffix *-ed* (e.g. *afford* → *afforded*) is extended to new verbs independently of their phonological properties. Also, in child language acquisition overregularization errors, i.e. irregular verbs inflected with the regular *-ed* form (e.g. *feel* → **feeled*, *sing* → **singed*), are a well-documented phenomenon (Marcus et al., 1992; Pinker, 1995b; Clahsen & Rothweiler, 1993; Maslen et al., 2004; Yang,

6 The learning and generalization of variable morphological rules: the case of the subjunctive

2002, 2016). This type of error sets in at a later time in language acquisition and it is understood to reflect the stage in which children have acquired the productive rule and are able to generalize beyond the input they receive (Pinker, 1995b). Importantly, evidence suggests that high-frequency irregular verb forms in adult speech are not regularized by children quite as often (Marcus et al., 1992). The dual-route model assumes that the derivation of regular forms takes place by the application of a “symbolic rule”. The symbolic rule is an operation over all the words that belong to a given syntactic category, regardless of their features (Say & Clahsen, 2002, p. 38). For example, in English, the rule that generates past tense forms is represented by the following operation: “add suffix *-ed* to X”. The variable X stands for any type of verb, regardless of its phonetic or semantic features. The lexical features of words are irrelevant to the application of the rule and so are their statistics of usage. Frequency of occurrence in speech is hypothesized to have no impact on the generalization and processing of regular morphological processes.

The symbolic rule is formally distinct from lexical representations. Regular complex forms do not need to be stored in the mental lexicon because they can be decomposed in constituent morphemes and derived in language production. The application of a symbolic rule is considered parsimonious in comparison to storing many complex forms and redundant context-dependent rules.

On the other hand, irregular forms are holistically stored in the mental lexicon. The knowledge of the irregular morphological pattern

6 The learning and generalization of variable morphological rules: the case of the subjunctive

is derived from the phonological association between the base and the complex form. The irregular pattern can be generalized only by analogy to existing words. This means that the similarity of the novel word to a set of learned instances induces the application of irregular morphology (Clahsen, 1999; Marslen-Wilson & Tyler, 1997; Pinker & Ullman, 2002). For instance, a hypothetical novel verb **thring* bears a phonological similarity with existing verbs *sing*, *ring*, *spring*. The past tense of these verbs is generated by a change of the stem vowel [ɪ]→[æ]: *sang*, *rang*, *sprang*. The similarity of the novel verb to this set may prompt speakers to adopt the same morphological pattern and produce the exceptional past tense form *thrang* (see Albright, 2002b, p. 1 for a similar example). This type of generalization is *lexically conditioned* (Bybee, 2002b). It means that it is influenced by the lexical features of the words that make up the speaker's lexical knowledge. Frequency of occurrence of the forms is only assumed to strengthen the irregular pattern.

At the other end of the spectrum, in analogy-based models like connectionist models (Rumelhart & McClelland, 1986; McClelland & Patterson, 2002; Hare et al., 1995; Skousen et al., 2002) and the network model (Bybee, 1995), rules and lexical representations are not formally distinct (Baayen, 1992, p. 138). Morphological knowledge is understood to arise by generalizing the patterns that emerge by relating together base words and complex word forms, which are initially all memorized. These relations established between all the base and complex words encountered are part of the speaker's lexical knowledge

6 The learning and generalization of variable morphological rules: the case of the subjunctive

and are called “networks of mappings” by Rumelhart and McClelland (1986) or “connections” by Bybee (1995). They are phonological and semantic connections between words. The number and types of shared phonological and semantic features are relevant in determining the solidity of the connections. When an unknown word is encountered, the higher its similarity to a set of learned items at the phonological and semantic level of representation, the higher the probability that the morphological pattern of that set will be applied (Hahn & Nakisa, 2000).

According to connectionist models and the network model, the statistics of previously learned input are crucial for the generalization of morphological patterns and the processing of complex forms. The two models, nonetheless, diverge in their predictions about how the statistics of usage impact productivity.

6.1.1 The role of frequency in analogy-based models

Connectionist models posit that both the size of the rule scope, i.e. the number of words that take the rule (type frequency), and the usage frequency of each form (token frequency) increase the productivity of the morphological pattern. For instance, the regular English past tense formed by adding suffix *-ed* (e.g. *stayed, talked*) to the verb base enjoys a very high type frequency: most English verbs take this morphological pattern while only approximately 200 verbs diverge from it. The rule’s very broad scope, which means that the *-ed* pattern will be encountered often, with a great variety of verbs (type frequency), and with no

6 The learning and generalization of variable morphological rules: the case of the subjunctive

particular restrictions, is predicted to increase the probability that speakers will use it to derive new forms (productivity). In comparison, let us consider a very small number of verbs derived from Germanic strong verbs. They are subject to stem *ablaut* [ɪ]→[æ] (e.g. *sing* → *sang*, *drink* → *drank*, *ring* → *rang*).

Connectionist models also predict that the frequent occurrence of one form with this type of morphological change in a high-frequency verb (token frequency), e.g. *began*, increases the productivity of the morphological pattern [ɪ]→[æ], even if this particular pattern cannot be found among many lexical items. On the other hand, the network model hypothesizes that only type frequency strengthens the connections between the words, on which the regularities are drawn, ultimately boosting the productivity of a morphological regularity. On the contrary, token frequency reinforces the lexical representation of the single word form and contributes to its lexical autonomy from the network (Bybee, 1995, pp. 433-435). In the case of the English past tense, high-frequency verbs like *take* [teɪk] → *took* [tʊk], *see* [si:] → *saw* [sɔ:] do not boost the productivity of the patterns they present [eɪ] → [ʊ], [i:] → [ɔ:]. Rather, token frequency is assumed to detract from the productivity of morphological rules because it weakens the connections between the base and the derived form in the lexical network.

According to Bybee (1988, 1995, pp. 432-433), this negative effect of token frequency explains the diachronic behavior of the English verbs. Indeed, it accounts for why high-frequency verbs resist

6 The learning and generalization of variable morphological rules: the case of the subjunctive

analogical levelling (i.e. the regular-*ed* past tense is still not possible with high-frequency verbs, e.g. *think* → **thinked*) but also for why they exhibit suppletive morphology in the first place. Complex forms that are highly frequent are simply stored independently of the other forms. Suppletion, this striking phonological difference between the base and the derived form, reflects the weakness of the connections.

Baayen (1992) points out that the network model “is ultimately based on the idea that productivity and regularity coincide”. He notes, however, that these notions are independent. In fact, “unproductive processes may be regular” (p. 140). He highlights the existence of productive processes which exhibit very low type frequencies, even lower than unproductive processes. Based on this observation he concludes that type frequency cannot be the only factor driving productivity.

Crucially, for both the network and the connectionist models there is no qualitative difference between regular and irregular morphological processes. Whether they are generalized to new forms, whether they resist analogical leveling, and how they are accessed in visual and auditory presentation only reflects the strength of the pattern on the basis of its type and token frequency of occurrence, and the number and types of shared features between novel items and existing sets.

The impact of type and token frequency and phonological similarity on the production of regular and irregular inflectional forms will be

6 The learning and generalization of variable morphological rules: the case of the subjunctive

explored in-depth in the next sections with the use of Italian verb morphology.

6.1.2 Summary

In the previous section three main theories of morphological learning and productivity were elucidated. The hypotheses derived from them are summarized here:

- a) Single-route rule-based model
No effects of frequency and similarity on the formation of complex forms
- b) Dual-mechanism model
Effects of frequency and similarity only on the formation of irregular forms
- c) Analogy-based models
Effects of frequency and similarity on the formation of both regular and irregular forms

In addition, it was shown that there is disagreement between two different analogy-based models concerning which frequency measure influences the productivity of morphological rules. Two main hypotheses have been put forward:

- I. Connectionist models
Both type and token frequency influence the generalization of morphological patterns
- II. The network model

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Only type frequency influences the generalization of morphological patterns

The experiment presented in section 6.5 tests the hypotheses made by the dual-mechanism model and the analogy-based models on the production and generalization of morphological variables. It also specifically examines the contradictory predictions of the network model and of connectionist models on the type of statistics upon which native speakers rely to build morphological knowledge and use it productively to construct new forms.

6.2 Italian verb morphology

In order to test the theoretical proposals presented in the previous section, a number of studies have focused on the verb system of the Romance languages, in particular on Italian verb morphology.

According to traditional linguistic treatments of Italian verbs, Italian, similar to other Romance languages, has three conjugation classes. They are distinguished by the thematic vowel, which appears between the verb root and the inflectional suffix in certain forms like the infinitive. 1st-conjugation verbs have the thematic vowel *-a-*, e.g. *mangi-are* “to eat”, 2nd-conjugations have thematic vowel *-e-*, e.g. *sorprend-ere* “to surprise”, and 3rd-conjugations have thematic vowel *-i-*, e.g. *soffr-ire* “to suffer”. Albright (2002a) observes that the 2nd conjugation can be further divided into two classes depending on

6 The learning and generalization of variable morphological rules: the case of the subjunctive

whether the stress falls on the root or on the suffix: *'sparg-ere* “to spread” vs. *tem-'ere* “to fear”.

Table 6.1 shows the relative distribution of these verb classes in Italian. It reports, along with mine, the frequency counts of the study that will be directly compared with the present experiment. The 1st conjugation is the largest class, grouping together the majority of verbs (74.9 %). It thus has by far the highest type frequency. Most of the verbs are regular, except for three verbs¹³. It is also variably considered the most productive (Albright, 2002a) or the only productive class in Italian (Dressler & Thornton, 1991, p. 5; Orsolini et al., 1998, p. 429): neologisms are formed with 1st-conjugation morphology, e.g. *googlare* “to google”, *bypassare* “to bypass”. Finally, it is the default class in the sense that its application is not restricted to verbs that show particular characteristics of any kind, e.g. semantic or phonological (Pinker & Prince, 1994; Bybee, 1995, pp. 438-439). It can select any type of verb root, variable Y in the following representation: “Y-are”. The 2nd and 3rd class have far fewer members, in the 2nd class they are mostly irregular (95% according to Say & Clahsen, 2002, p. 96). Additionally, most high-frequency verbs belong to the 2nd conjugation (infinitive suffix *-ere/-ére*). They are largely considered to be no longer productive

¹³ One of the three irregular verbs, *fare* “do/make”, derives from Latin *facēre* and is best understood as part of the 2nd conjugation because it takes vowel -e in the tensed forms that exhibit the old Latin root, e.g. 1st pers. sg. imperfect indicative *facevo*, and imperfect subjunctive *facessi* (Say & Clahsen, 2002, p. 96).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

in Italian¹⁴. For these properties, the Italian conjugation classes have been used in the literature like the English past tense to explore whether there are fundamental differences in the way regular and irregular verbs are produced and parsed. In particular, the 1st conjugation is considered to have regular and productive morphology (more or less equivalent to the past tense *-ed*) whereas the morphological patterns of the 2nd and 3rd conjugations are considered irregular and unproductive (like the various irregular verb classes in English, e.g. *begin* > *began*, *break* > *broke*). The following hypotheses have been tested:

1) Dual-mechanism model

Generalization of regular affixal morphology of the productive class (1st conjugation) takes place by application of a context-free symbolic rule and is insensitive to phonological properties of the existing verbs.

Novel use of irregular unproductive morphology (2nd and 3rd conjugations) is rare and prompted by phonological similarity to attested examples. (Say & Clahsen, 2002; Veríssimo & Clahsen, 2014).

2) Analogy-based accounts

Graded generalization of all inflectional paradigms is based on phonological similarity to attested examples and sensitive to usage statistics of previously learned input. (Eddington, 2002; Colombo et al., 2006)

¹⁴ The 3rd conjugation *-ire* was used to create inchoative verbs out of adjectives (e.g. adj. *giallo* “yellow” > *ingiallire* “to become yellow”).

6 *The learning and generalization of variable morphological rules: the case of the subjunctive*

Table 6.1

Relative distribution of verbs by conjugation classes in Italian

	sample: 6897 verbs (Olivetti Italian Dictionary)	sample: 3219 verbs (itWaC, Freq. 1*Million)	sample: 2022 verbs (De Mauro et al. 1993 in Albright, 2002a)
1 st conjugation	78.4%	74.9%	72.4%
2 nd conjugation	10.8%	13.2%	16%
3 rd conjugation	10.7%	10.3%	9.7%
-ére verbs	--	1,1%	2,1%

According to Albright (2002a, p. 691) and Orsolini and Marslen-Wilson (1997), Italian is better suited than English for exploring effects of frequency and similarity. In fact, in English there is a far greater imbalance between the regular verbs and the irregular verbs. The latter are only approximately 160 in total and are split into many different classes (Bybee & Slobin, 1982; Pinker & Prince, 1988). Each class has extremely low type frequencies: most of them do not have more than 6-7 verbs. Possible similarity effects on the formation of new words based on type frequency would be very weak. The regular class is remarkably

6 *The learning and generalization of variable morphological rules: the case of the subjunctive*

much larger and groups together extremely diverse verb types. It is therefore difficult to find clear effects based on the statistics of the lexicon and discriminate between gradient similarity and an unbound similarity-insensitive mechanism. On the other hand, the unproductive classes in Italian, some of which gather most irregular verbs, are “potentially large enough to attract new words” (Albright, 2002a, p. 691). The productive pattern, encompassing fewer verbs, offers more possibility to discover lexical effects.

To these observations I add that a major advantage posed by the Italian subjunctive or the infinitive is that it eliminates the strong correlation between token frequency and non-compositionality. The English irregular past tense forms are derived through changes of the stem that are often unpredictable, e.g. *swim* > *swam*, *bite* > *bit*, *write* > *wrote*. Crucially, the resulting forms cannot be segmented into lexical affix and inflectional affix. In Italian, even highly suppletive verbs, e.g. *andare* “to go”, *potere* “can” that radically change the stem, when inflected, take an affix that conveys the relevant grammatical information. The affix is highly predictable. The resulting form can be segmented into lexical morpheme and grammatical morpheme, e.g. irregular verb *andare* “go” (infinitive) > *vad-ano* (subjunctive) vs. regular verb *cedere* “yield” (infinitive) > *ced-ano* (subjunctive). In both cases the same compositional morpheme *-ano* conveys the grammatical information: third person plural subjunctive present. Therefore, high frequency verbs in non-productive classes still exhibit a segmentable pattern.

6.3 Psycholinguistic studies on Italian verb morphology

The studies on Romance verb morphology have found contradicting results so far (Caramazza et al., 1988; Cappa & Ullman, 1998; Orsolini & Marslen-Wilson, 1997; Orsolini et al., 1998; Say & Clahsen, 2002; Albright, 2002a, 2002b; Veríssimo & Clahsen, 2014 among others). Orsolini and Marslen-Wilson (1997) and Orsolini et al. (1998) studied regular and irregular past forms in Italian. In a priming experiment, Orsolini & Marslen-Wilson (1997) observed that both regular inflected pairs, e.g. *giocarono/giocare* (1st conjugation), and irregular inflected pairs, e.g. *scesero/scendere* (2nd conjugation), generated the same amount of priming. The results, according to the authors, support the hypothesis that both regular and irregular forms are decomposed into constituent morphemes. In a second experiment, where subjects were asked to produce past tense forms for novel verbs, they found that phonological similarity influenced the production of both regular and irregular patterns, an effect predicted by analogy-based models which posit gradient morphological generalizations based on analogy. Orsolini et al. (1998) investigated the spontaneous production of past tense forms in children aged between 4 and 10. They found that the 1st conjugation pattern was not overgeneralized with respect to the non-productive pattern of the 2nd conjugation. In fact, the opposite happened. The evidence disfavors the dual-mechanism model, which assumes that errors of overgeneralization are due to the fact that regular processes are produced by application of a context-insensitive rule. Shortcomings of the experimental designs are discussed in Say &

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Clahsen (2002, pp. 105-107) who also present a reanalysis of the experiment results. According to Say and Clahsen, a dual-mechanism model best captures the linguistic behavior elicited in their experiment. They had adult participants produce past participle forms for nonce verbs in a written production task. Some of the stimuli bore similarities to existing regular and irregular verbs of the 2nd and 3rd conjugation. Similarity was based on rhyme: the stimuli were constructed so as to share the final vowel and consonant(s) of existing verb roots. Another group of stimuli did not share this particular phonological segment with any of the existing verbs (no similarity condition). They found that the 1st conjugation morphology (the regular and productive pattern) was generalized to verbs that are similar to other conjugations 57% of the times. Whereas the non-default patterns of the 2nd and 3rd conjugations overgeneralized only 3% or 4% of the times, a statistically significant difference. However, they did not include stimuli that are similar to 1st conjugation verbs as an experimental condition. Therefore, the study could not establish whether similarity influences the production of the default regular pattern. Also, the experiment showed ambiguous results with regard to the role of frequency. Contrary to the expectations of the researchers, the irregular verbs of the 2nd conjugation did not attract more members as a function of their frequency. However, 2nd conjugation regular verbs and both regular and irregular 3rd conjugation verbs did.

In that experiment, similarity was operationalized as a rigid concept by taking into account the verb rhyme only. This did not allow to test

6 The learning and generalization of variable morphological rules: the case of the subjunctive

effects of gradient similarity and might have generated possible confounds. For instance, verb roots ending in segment /ɛm/ have a very high probability (96%) of taking a 2nd conjugation infinitive morpheme -ere (e.g. *fremere* “to tingle”, *premere* “to push”). However, verb roots ending in phoneme /m/ have a high probability (81%) of taking the 1st conjugation morphology (e.g. *amare* “to love”, *animare* “to animate”). Also, the two groups largely differ in size (type frequency), the second one being much larger than the first. This difference is relevant in models that assume type frequency to be a major predictor of rule generalization (Bybee, 1995 among others).

The Minimal Generalization Learner (MGL), a computational model developed by Albright & Hayes (2003) (also Albright, 2009) offers a more systematic and refined way of appraising the role of gradient phonological similarity and frequency on the production of novel forms. It was used by Albright (2002a) to test the acceptability rate of novel infinitives in Italian and by Veríssimo & Clahsen (2014) to establish whether the production of Portuguese infinitives is probabilistic and driven by similarity. The experiment I present is based on these studies and relies on the MGL for the construction of the experiment materials. The characteristics of the model are presented in the following section along with a review of the studies on Romance verb morphology based on the MGL.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Table 6.2
A representation of the phonological environments discovered by the MGL

	Residue	Shared features	Shared segments	Change
Comparing				
A		<i>f</i>	<i>rem</i>	<i>o</i> → <i>ere</i>
with B	<i>s</i>	<i>p</i>	<i>rem</i>	<i>o</i> → <i>ere</i>
		[+consonant		
Yields C	X	- voiced	<i>rem</i>	<i>o</i> → <i>ere</i>
		- dental		
		etc]		

Note. Adapted from *Islands of reliability for regular morphology: Evidence from Italian* by Albright (2002a, p. 687). Copyright 2002 by Language.

Table 6.2 shows how the pairs are evaluated by the model to form a rule that derives the infinitive out of the 1st sg. present indicative. The two verb forms undergo the same change, suffix *-o* is substituted by infinitive suffix *-ere* (*o* → *ere*). Additionally, they share a particular phonological environment to the immediate left of the change (C__#). This is made up of the phonological segment /*rem*/ and, next to it, the phonological features: voiceless consonant. Finally, the MGL

6 The learning and generalization of variable morphological rules: the case of the subjunctive

recognizes a residue (represented by X) comprised of the phonological material that the pairs do not have in common.

Rule in SPE notation:

$$o \rightarrow ere/ [X[+cont, -voicd] rem __]$$

The rule that emerges from this pairwise comparison is then evaluated against all the words available in the lexicon to find out how many words share the phonetic environment (the scope of the rule or the rule's type frequency) and how many of them actually participate in the morphological change (hits).

The ratio of the hits to the scope gives a measure of the reliability (or confidence) of the rule. It reflects the probability that the rule will be applied in that particular context. If the rule has many exceptions, its reliability will be accordingly low.

1. Definition of rule reliability based on type frequency in Albright (2002a, p. 687):

$$\frac{\text{Number of forms included in the rule's structural change (its hits)}}{\text{Number of forms included in the rule's structural description (its scope)}}$$

For instance, in Italian, out of 213 verb stems ending in the phoneme /m/, 181 types take infinitive suffix -are. This gives a rule reliability of $181/213 = .84$. This score is then adjusted using lower confidence limit statistics (at a confidence level of 75%) based on Mikheev (1997). The

6 The learning and generalization of variable morphological rules: the case of the subjunctive

effect of this adjustment is to downgrade the reliability of rules that are based on very few words.

As Albright notes, this particular definition of rule reliability is based on type frequency and is meant to capture the hypothesis according to which type frequency is the most important factor in the generalization and production of morphological patterns (Bybee, 1995): in order to discover a rule and apply it productively to new words, there must be evidence of it in a sufficient number of different words in the linguistic input. Token frequency, i.e. the occurrences of the morphological pattern in a given form, is irrelevant.

As was previously discussed, connectionist models (Rumelhart & McClelland, 1986; McClelland & Patterson, 2002) postulate that token frequency also plays a fundamental role. In order to enable the comparison between the different theoretical proposals, the model additionally allows to calculate the reliability of the morphological pattern based solely on token frequency (2.) and, crucially, on a combined measure of type and token frequency (3.).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

2. Rule reliability based on token frequency

$$\frac{\text{Frequency of forms included in the rule's structural change (hits)}}{\text{Frequency of forms included in the rule's structural description (scope)}}$$

3. Rule reliability based on a combined measure of type and token frequency

$$\frac{\text{Hit types}}{\text{Scope types}} \times \frac{\text{Hit tokens}}{\text{Scope tokens}}$$

Beyond all the phonologically sensitive rules, the model also discovers general context-free rules, e.g. rule to derive the infinitive from the 1sg. present: (*o* → *are* / X___), where X represents any type of phonetic sequence. They are formed when the lexical items that participate in a given morphological change have no phonological material in common. The context-free rule is the equivalent of the symbolic rule theorized by the dual-mechanism model (see again Say & Clahsen, 2002) and applies to any type of verb. In our case, the rule simply states “change suffix *-o* into *-are*”.

However, the MGL differs from the dual-mechanism model in a crucial way. It retains all the context-sensitive rules even when the context-free rule is discovered. Ultimately, only the rule with the highest reliability is chosen by the model to derive the new word form. The context-sensitive rules will be preferred to the symbolic rule if they have a higher reliability. The model relies on the statistics and the lexical characteristics of the input in all instances to form its grammar.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

On the contrary, dual-mechanism models as proposed by Say & Clahsen (2002) and Veríssimo & Clahsen (2014) assume that context-sensitive rules will be stored only for irregular unproductive patterns (e.g. infinitives *-ere*, *-ére*, and *-ire*). The regular rule (e.g. infinitive *-are*) will always be applied to novel words independently of their phonological shape.

6.4.1 Results of the experiments based on the MGL algorithm

Albright (2002a) tested these predictions on Italian infinitives using an acceptability judgement task. Adult participants were asked to rate the well-formedness of infinitive forms of novel verbs on a 7-point scale. The stimuli were constructed using the MGL and exhibited various phonological environments with various reliability scores for each conjugation.

The study found that acceptability ratings of novel infinitives belonging to all the conjugations significantly correlated with the MGL rule reliabilities. Importantly, the acceptability rates of the regular default pattern were influenced by similarity to existing words and by the statistics of the lexicon just as for the irregular patterns. Such evidence contradicts the predictions of the dual-mechanism model and lends support to analogy-based models. The study additionally tested the performance of the MGL against human performance with the rule reliability scores based on type frequency and on a combined measure of type and token frequency. Albright reports that all the measures

6 The learning and generalization of variable morphological rules: the case of the subjunctive

perform similarly. Thus, no definitive answer could be given on the diverging predictions of the network model and of connectionist models. According to the author, the experiment was not constructed to tease apart the effects of type and token frequency, something that will be carefully considered in the experiment reported in the present dissertation.

Verissimo & Clahsen (2014) conducted an analogous study on Portuguese conjugation classes using the MGL computational model to establish whether similar results are obtained for the formation of Portuguese infinitives.

Portuguese verb morphology and the distribution of verbs among conjugation classes is very similar to Italian. The 1st conjugation with infinitive suffix -ar (e.g. *amar* “to love”) is the largest and most productive class. Based on a corpus of Portuguese (Bacelar do Nascimento et al., 2000), the authors report that 3,396 verbs belong to the 1st conjugation, and in comparison only 380 types are grouped in the 2nd conjugation, and 348 in the 3rd. In addition, most low frequency verbs belong to the 1st class, a statistic that is assumed to boost rule productivity (Baayen, 1992, p. 140).

Table 6.1 above shows a direct comparison between the size of the verb classes in Italian and Portuguese. It is reported again in Table 6.3. The 1st class in Portuguese, based on the authors’ count, is even larger in comparison to Italian. Its power of attraction based on its type frequency should prove to be even greater. It is also identified as the default class. It has “unrestricted productivity” and novel words that are

6 The learning and generalization of variable morphological rules: the case of the subjunctive

introduced into the language, such as foreign borrowings, are assigned to the 1st class (Veríssimo & Clahsen, 2014, p. 63).

Table 6.3
Relative distribution of the verb classes in Italian and Portuguese

		Italian	Portuguese
		sample:	sample: 4126
		2022	verbs
		(De Mauro	(Bacelar do
		et al. 1993	Nascimento et
		in	al. 2000 in
		Albright,	Veríssimo and
		2002)	Clahsen, 2014)
1 st conjugation	74.9%	72.4%	82.3%
2 nd conjugation	13.2%	16%	9.2%
3 rd conjugation	10.3%	9.7%	8.4%

The authors conducted a written production task. Adult participants were asked to write the appropriate form of novel verbs in a sentence. The sentence required the infinitive in the test condition. Following Albright (2002a), the verbs were first presented in the 1st person singular present (indicative) which has suffix *-o* and is ambiguous with regard to conjugation class. The participants had to transform the verb

6 The learning and generalization of variable morphological rules: the case of the subjunctive

form to use an infinitive that would appropriately fit the test sentence. They were thus faced with choosing one of the infinitive affixes of the three conjugations. See the following example provided in Veríssimo and Clahsen (2014, p. 67):

Quase sempre acuo sozinho.

Mas amanhã vou _____ acompanhado.

‘I almost always acuo alone.’

‘But tomorrow I will _____ with someone.’

The novel verbs, 78 in total, were created so as to fall within different “island rules” that represent the whole spectrum of reliabilities given by the MGL: from high to low reliabilities for each conjugation. The authors found that similarity to existing verbs only positively influenced production of novel infinitives for the 2nd and 3rd conjugation (irregulars). This was not the case for the 1st conjugation, where the production of the regular pattern is insensitive to the phonological characteristics and the statistics of the lexicon. For the 1st conjugation, there was only a negative effect: similarity to other classes significantly decreased the probability of using 1st conjugation suffix *-ar*. They concluded that the results cannot be captured by a single analogy-driven mechanism but that they are in line with the predictions of the dual-mechanism model.

Clearly the results displayed by Albright (2002a) and Veríssimo & Clahsen (2014) are at odds. An obvious observation is that the two

6 The learning and generalization of variable morphological rules: the case of the subjunctive

experimental tasks differ, one is an acceptability judgment task, whereas the other is a written production task. The results may reflect the distinct mechanisms or strategies adopted by the participants to assess the well-formedness of new infinitives on the one hand and to produce them on the other.

Looking closely at the two studies, however, a crucial difference arises. The MGL model discovers a default context-free rule for both languages. This represents the lower boundary: more specific rules, which specify a phonological context, will only be considered only if they show a higher reliability value than the context-free rule.

The context-free rule for the 1st conjugation in Italian (*o* → *are/X*____) has a very high adjusted reliability, .717, as reported by Albright. This is in line with the size of the conjugation classes reported in his study. In the language sample used by Albright 72.4% of the Italian verbs take the regular infinitive suffix *-are*, which is 1,461 out of 2,022 verbs (see Table 6.3 above). In Portuguese, however, the context-free rule is reported to have a much lower reliability value (Verissimo & Clahsen, 2014, p. 67). The number is not explicitly stated in the article but it should amount to .459 or lower (this is the lowest reliability value reported for the 1st conjugation on the list of experiment stimuli reported in the Appendix on p. 77. Overall, 31 experimental items have a reliability value below .584). At first glance, the difference seems puzzling. Italian and Portuguese have comparable class sizes and, similar to Italian, the Portuguese 1st conjugation displays a regular rule to derive the infinitive from the 1sg. present in most verbs (*o*→

6 The learning and generalization of variable morphological rules: the case of the subjunctive

ar/X__) (Veríssimo & Clahsen, 2014, p. 63). In the sample used by the authors, the 1st class is even larger in comparison to the other classes, 82.3% of verbs follow this pattern (shown in Table 6.3), totaling 3,396 out of 4,124 (reported on p. 63). This should give an unadjusted context-free rule reliability of $3,396/4,124 = .823$. The authors only used the infinitive forms that actually appear in the corpus. Based on the numbers reported, one must conclude that the proportions of infinitives of each conjugation class greatly diverge from those found in Italian (see again Table 5.3). Beyond this observation, the striking difference between the values reported in the two studies cannot be explained at present. The studies rely on the MGL reliabilities to assess the effects of frequency and phonological similarity. The difference must be addressed in order to properly evaluate the results. In particular, the low reliabilities for the 1st conjugation, if incorrect, could account for why the MGL does not correlate with human behavior in this category. An analysis of Portuguese verb morphology and its phonotactics is beyond the scope of the present dissertation and will be left to a future endeavor.

In the next section I present an experiment I conducted on the production of Italian infinitives and subjunctives. The experiment aims at addressing the diverging results of the two previous studies. In particular, I return to Italian verb morphology to investigate the production of novel infinitives and subjunctives. Using a task similar to the one presented in Veríssimo & Clahsen (2014), it is my goal to find out whether the behavior observed by Albright on Italian is replicated.

6.5 Experiment on the production of novel infinitives and subjunctives: premises and hypotheses

The experiment investigates the formation of novel infinitives and novel subjunctives in Italian. The infinitive has a different suffix for each conjugation: *-are*, *-ere*, *é-re*, *-ire*. Instead, the present subjunctive exhibits one set of suffixes for the 1st conjugation and another set of suffixes for all the others¹⁶ (see Table 6.4). This generates a binary distinction between the productive, default class and the unproductive, non-default classes. In the first person singular form of the present elicited in the experiment, morpheme *-i* is selected for verbs of the 1st conjugation and morpheme *-a* for all others.

Table 6.4
Inflectional paradigm of the Italian subjunctive present for each conjugation

	1 st	2 nd	3 rd	4 th
Verb	conjugation	conjugation	conjugation	conjugation
person	<i>am-a-re</i>	<i>ricev-e-re</i>	<i>accad-é-re</i>	<i>dorm-i-re</i>
	“love”	“receive”	“happen”	“sleep”
1/2/3				
SG	<i>am-i</i>	<i>ricev-a</i>	<i>accad-a</i>	<i>dorm-a</i>
3 PL	<i>am-ino</i>	<i>ricev-ano</i>	<i>accad-ano</i>	<i>dorm-ano</i>

¹⁶ The 1st pl. and the 2nd pl. forms of the subjunctive present are syncretic. They exhibit the same suffix (respectively *-iamo*, *-iate*) for all the conjugations.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

The following hypotheses derived from analogy-based models (Eddington, 2002; Colombo et al., 2006) and the dual-mechanism model (Say & Clahsen, 2002; Veríssimo & Clahsen, 2014) are tested:

- (1) The production of all complex forms of the four conjugations is based on similarity and sensitive to the usage statistics of previously learned input (analogy-based models).
- (2) Only the morphological patterns of the 2nd and 3rd and 4th conjugation (irregular pattern) are generalized through analogy, while the application of 1st conjugation morphology (regular pattern) is insensitive to the phonological properties and the statistics of existing verbs (dual-mechanism model).

Concerning hypothesis (1), the experiment explicitly tests which type of statistics, if any, influence the generalization of a morphological pattern:

- (1a) Only type frequency boosts the productivity of a morphological pattern and increases the probability that it will be generalized to new words. Token frequency negatively influences the productivity of the morphological schema (Bybee's network model).
- (1b) Both type and token frequency positively affect the probability that a morphological pattern will be extended to novel words (connectionist models and Baayen et al. 2002)

The hypotheses are tested using the MGL algorithm (Albright, 2002a; Albright & Hayes, 2003; Veríssimo & Clahsen, 2014 among

6 The learning and generalization of variable morphological rules: the case of the subjunctive

others). As was discussed above, the rules derived by the MGL are based on phonological similarity to existing words and are probabilistic. The MGL rule reliabilities reflect the probability that a rule will be selected based on the frequency of the morphological pattern among word types (type frequency) or its absolute frequency of occurrence in speech (both type and token frequency). With its architecture, the MGL captures the hypotheses of analogy-based models (see again section 6). The experiment tests whether the MGL rule reliabilities correlate with human responses. In order to do that, regression analysis is used.

According to the predictions of analogy-based models (Albright, 2002a), the MGL rule reliabilities should correlate with the production of all complex forms, both regular (1st conjugation responses) and irregular (2nd /3rd /4th conjugation responses). Thus, a positive linear correlation is expected in all cases: the higher the rule reliability, the higher the probability that speakers will use that rule to derive the complex form.

On the other hand, according to the predictions of the dual-mechanism model as formulated by Veríssimo & Clahsen (2014), the MGL rule reliabilities should only positively correlate with the production of infinitives and subjunctives of the 2nd /3rd /4th conjugations (irregular morphological patterns). No correlation should be established between the regular pattern (1st-conjugation infinitives and subjunctives) and the MGL's predicted responses. The reason for the lack of a correlation is that, according to the model, regular forms

6 The learning and generalization of variable morphological rules: the case of the subjunctive

are produced exclusively by a symbolic context-free rule, e.g infinitive X-*are*. The application of the rule is not probabilistic: neither similarity to existing regulars nor frequency influences the probability that the regular morphological pattern is selected.

The opposing hypotheses on how the regular 1st conjugation pattern is extended to new verbs are illustrated using two novel verb forms: *spurimo* and *aremizzo*¹⁷ (1sg. present indicative). For verb form *spurimo*, according to the MGL model, there is a 79% adjusted probability that speakers will use rule ($o \rightarrow \text{are/m}___$) to derive the infinitive form *spurimare*. The probability reflects the fact that 91 out of 111 verbs whose stem ends with phoneme /m/ take regular infinitive suffix -*are* in the corpus. On the other hand, for novel verb *aremizzo* there is a 99% adjusted probability that speakers will use rule ($o \rightarrow \text{are/z}___$): 143 out of 143 of verbs whose stem ends in phoneme /dz/ take suffix -*are*. According to analogy-based models, the probability that experiment participants select suffix -*are* is greater for verb *aremizzo* than it is for *spurimo*. The experiment responses should reflect this prediction.

According to the dual-mechanism model, both verbs, which bear no similarity to existing irregular forms, are derived by a context-insensitive rule ($o \rightarrow \text{are/X}$) regardless of their similarity to existing regulars. This derivation is not probabilistic, thus the responses should

¹⁷ The two pseudoverbs are part of the experiment stimuli. They are presented in the 1sg. present indicative which displays the same suffix -*o* in all conjugations.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

be equal for the two verbs and the probability of application of the regular pattern *-are* should both be 1.

As already hinted in the introduction to this chapter, no study exploring the above-mentioned hypotheses has been conducted on the subjunctive. Also, studies investigating the use of the subjunctive have always attempted to give an overall measure of its productivity. No other work has analyzed the productivity rates of the subjunctive paradigm of each conjugation.

Comparing the subjunctive and the infinitive can shed light on how usage statistics influences adults' grammar. As a matter of fact, the statistics of the subjunctive and the infinitive in speech differ in some crucial aspects. Table 5.3, Table 5.4 and Table 5.2 above report relevant descriptive measures of the subjunctive and infinitive gathered from the corpus study. They are displayed again in Table 6.5, Table 6.6 and Table 6.7 respectively for ease of consultation.

Table 6.5

Type/token distribution of infinitives in C-ORAL-ROM

Infinitives	Verb types	Verb types in %	Verb tokens	Type/token ratio
overall	1009	100%	8925	0.11
1 st conjugation	689	68.2%	4492	0.15
2 nd conjugation	191	18.9%	2592	0.07
3 rd conjugation	129	12.7%	1752	0.07

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Table 6.6

Type/token distribution of subjunctives in C-ORAL-ROM

Subjunctives	Verb types	Verb types in %	Verb tokens	Type/token ratio
overall	156	100%	1148	0.13
1 st conjugation	83	53%	276	0.30
2 nd conjugation	37	23.7%	710	0.05
3 rd conjugation	36	23%	162	0.22

Table 6.7

Distribution of subjunctive and infinitive morphology among high-frequency verbs in C-ORAL-ROM

	Verbs	Subjunctive	Infinitive
<i>essere</i>	‘be’	34.8%	6.8%
<i>avere</i>	‘have’	10.2%	2.4%
<i>potere</i>	‘can’	7.1%	0.3%
<i>fare</i>	‘do’	5.1%	11.0%
<i>venire</i>	‘come’	3.6%	0.7%
<i>sentire</i>	‘listen’	3.0%	0.7%
<i>dire</i>	‘say’	3.0%	7.7%

6 The learning and generalization of variable morphological rules: the case of the subjunctive

The infinitive forms of the 1st conjugation (the largest class) occur most frequently in speech (4492 infinitive tokens) and with the highest number of different verb types (68.2% of all the infinitives). The infinitive suffix of the other conjugations is far less represented in speech in both types and tokens (see Table 6.5). Thus, there is a large imbalance between the frequency of the regular pattern and the frequency of the the irregular patterns.

On the other hand, the distribution of subjunctive forms among the conjugations differs considerably. The 1st conjugation exhibits only slightly more subjunctive types than the other conjugations (83 types vs 73 types of the other conjugations¹⁸). Additionally, most subjunctive tokens occur with verbs of the 2nd conjugation (710 tokens in total vs. 276 subjunctive tokens of the 1st conjugationclass). The reason lies in the fact that the majority of subjunctives (66%) occur with high-frequency verbs (see Table 6.7). Most of them belong to the 2nd and 3rd conjugation (-*ere*, -*ére*, -*ire*), e.g. *si-a* “be”, *abbi-a* “have”, *poss-a* “can”, *facci-a* “do”. Thus, the non-default irregular subjunctive pattern has higher type frequency than the non-default irregular infinitive patterns. It also has an exceptionally high token frequency.

If type and token frequency boost the productivity of a morphological pattern, as hypothesized in 1b, the non-default

¹⁸ The 2nd conjugation (37 verb types) and 3rd conjugation (36 verb types) are considered together because they exhibit the same subjunctive pattern. On the other hand, the infinitive has a different morphological pattern for each conjugation.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

subjunctive pattern will have a greater chance of being selected with new words than the non-default infinitive patterns.

6.5.1 Method

6.5.1.1 Participants

Thirty-five adult native speakers of Italian aged between twenty and fifty-one (19 females) participated in an oral production task. They were paid for their participation. They had all completed at least 13 years of school education and had normal or corrected-to-normal vision. The goal of the experiment was not disclosed to them. Before starting the experiment, all participants signed a consent agreement and were instructed that they could withdraw their consent at any time during the experiment and after.

6.5.1.2 Procedure

The experiment was programmed with PsychoPy2 v1.84.2 (Peirce, 2007; Peirce et al., 2019) and conducted with the use of a MacBook Pro laptop. It was presented to participants on a 21.5 inch monitor (Acer K2-Series). The participants wore headphones with a microphone connected to the audio mixer “Focusrite iTrack Solo” in order to get echo feedback. They interacted with the experiment interface through a MacBook Bluetooth keyboard. Their oral responses were recorded with the *microphone* module of Psychopy. They were later transcribed and

6 The learning and generalization of variable morphological rules: the case of the subjunctive

annotated in order to conduct the statistical analyses. The participants received the following instructions on the monitor¹⁹:

“The experiment presents sentence pairs. One sentence pair at a time appears on the screen.

The first sentence above contains a word highlighted in bold. This word does not exist in the Italian language (new word). The second sentence below contains a blank space. Fill in the blank space with the new word. Transform the word appropriately so that it fits in the context of the second sentence.”

Each novel verb was presented only once in the experimental condition. As the instructions suggest, it was embedded in a sentence (input sentence). Participants were instructed to transform the new verb in order to fill in the blank space in the sentence underneath (test sentence).

The test sentence in the experimental condition required either a subjunctive or an infinitive. The subjunctive sentences always required a 1sg. subjunctive present form. Following Say & Clahsen (2002) and Veríssimo & Clahsen (2014), the novel verb in the input sentence was always presented in the 1sg. present indicative with suffix *-o*. This form is ambiguous for conjugation. In order to produce the appropriate infinitive or subjunctive form, participants had to implicitly assign the nonce verb to a conjugation class. For instance, in example (A) participants can produce one among the possible infinitive forms: e.g.

¹⁹ The instructions were presented in Italian. An English translation has been provided.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

burt-are (1st conjugation), *burt-ere* or *burt-ére* (2nd conjugation), *burt-ire* (3rd conjugation). In example (B) the appropriate response is a subjunctive: e.g. *trud-i* (1st-conjugation subjunctive) or *trud-a* (all other conjugations). The participants were told that there was no right or wrong answer and were instructed to respond instinctively.

(A)

(Input sentence) Non **burto** nemmeno io.

‘I don’t **burto** either.’

(Test sentence) Dovrebbe potere _____ anche lui. (**infinitive**)

‘He should be able to _____ as well.’

(B)

(Input sentence) Oggi **trudo** in silenzio.

‘Today I **trudo** silently.’

(Test sentence) Bisogna che io _____ in silenzio. (**subjunctive**)

‘It is necessary that I _____ silently.’

The participants could read the sentence pairs at their own pace. Once they were ready to give their response, they tapped the spacebar on the keyboard. After hearing a beep sound in the headphones, they pronounced the appropriate word aloud. The responses were thus elicited orally as opposed to the previous experiments which had participants write their answer on paper. This was done for two reasons. As it was discussed, Italian infinitive forms ending in *-ere* can either have a stress on the verb root or on the suffix, e.g. *'sparg-ere* “to spread”

6 The learning and generalization of variable morphological rules: the case of the subjunctive

vs. *tem-'ere* “to fear”. Also, subjunctive forms of more than two syllables can vary in stress. The stress however is not marked in writing. It would have been impossible to recognize where the participants had intended to place the stress and what type of word they had intended to produce. Additionally, an oral response was deemed less calculated and more instinctive. After the response was recorded, the experiment automatically presented the next sentence pair. Only one sentence pair at a time was displayed on the screen. After giving their answer, the participants could not come back and change it.

The participants first read the experiment instructions and had the chance to ask questions to the researcher. The experiment started after completion of a practice phase with 4 practice sentence pairs. Upon completion of the practice phase, the experimenter made sure that the participants had understood the instructions and were ready to start. After half of the experiment was completed (after 45 stimuli), the participants had a break. The mandatory break lasted 3 minutes, during which time the monitor froze and the participants could not interact with it. When the scheduled break time was up, the participants could choose to have more time to rest. When they were ready to restart the experiment, they tapped the spacebar. The break was designed to limit the level of fatigue. After completion, the participants reported that they did not experience fatigue during the task. The average experiment duration was 35 minutes.

6.5.1.3 Simulation with the MGL model

The experiment materials were constructed using the MGL model (Albright & Hayes, 2003) following Albright (2002a) and Veríssimo & Clahsen (2014). The lexical sample from which the model learned the morphological rules was made of 3219 verbs, i.e. the verbs that occur at least one time in a million words in the ItWac corpus (Baroni et al., 2009).

The MGL takes pairs of words in a list as its input. In our case two lists were created. One with 1sg. present indicative - infinitive pairs (e.g. *parlo* → *parlare* “I talk” → “to talk”) and another with 1sg. present indicative – 1sg. present subjunctive forms (e.g. *dormo* “I sleep” → *dorma* “I sleep.SUBJ”). They were entered into the model separately. Thus, the MGL first learned the rules to derive infinitives and then those to derive the subjunctives. The two lists of form pairs were transcribed phonetically so that the MGL could recognize the phonetic environment of each rule. The model assigns a confidence score to each rule, the rule reliability. Three types of rule reliabilities were calculated according to the methods described above:

- 1) rule reliabilities based on type frequency
- 2) rule reliabilities based on token frequency
- 3) rule reliabilities based on a combined measure of type and token frequency

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Token frequencies of each verb were extracted from the itWac (Baroni et al., 2009), a corpus big enough to differentiate low-frequency verbs.

The rules were adjusted using a lower confidence limit of 75% following Albright (2002a) and Verissimo & Clahsen (2014).

For the creation of the stimuli and the statistical analysis the verb classes were grouped as follows: 1st conjugation gathers the verbs whose infinitive form ends in *-are* (e.g. *giocare* “to play”), 2nd conjugation was assigned to the verbs ending in *-ere* with the stress on the root (e.g. *correre* /’korrere/ “to run”), 3rd conjugation was assigned to the verbs ending in *-ére* with a stress on the suffix (e.g. *temere* /te’mere/ “to fear). Finally, the 4th class collects verbs which take suffix *-ire* (e.g. *scoprire* “to discover”).²⁰

6.5.1.4 Materials

The MGL rules with their environments and reliability scores served as the basis to form the novel verbs presented in the experiment. 86 novel verbs were created. They were designed so as to represent the entire spectrum of rule reliabilities discovered by the MGL for each conjugation, from high to low, thereby increasing the statistical power to detect continuous effects of frequency and similarity (see Appendix

²⁰ In traditional grammars, the Italian conjugation classes are considered to consist of three groups and are based on the theme vowel. Following Albright (2002a), the verbs that are usually grouped in the 2nd conjugation with theme vowel *-e-* were further divided in two groups based on the word stress of their infinitive form. As a result, I obtained four classes.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

E). For the 1st conjugation, the MGL also discovered a context-free rule. This had a high adjusted reliability of .754, a number similar to the one reported in Albright (2002a, p. 717). The reliability scores reflect the percentage of verbs that take the 1st conjugation out of all the verbs in the respective samples: 74.9% in my sample and 72.4% in Albright's sample (see again Table 6.3).

The reliability of the context-free rule represents the lower boundary and necessarily limits the range and variance of the experiment stimuli for the 1st conjugation (see Table 6.8 for the variance, range and standard deviation of the experiment stimuli).

Among the experiment items, there are forms whose most reliable rule outputs the morphological change of the non-default conjugations (infinitive suffixes *-ere*, *-ére*, *-ire* with or without changes of the stem, and subjunctive *-a* with or without stem changes).

The experiment items were explicitly designed to tease apart effects of type and token frequency. This was achieved by creating items whose reliability diverges as much as possible when calculated according to type frequency and to token frequency; at least a 0.10 point difference but in many cases the difference exceeds 0.30. At least 20 items were modeled on existing words with exceptional patterns and high token frequency. The MGL thus outputs rules that operate phonological changes of the stem in addition to adding the suffix. An example is given in Table 6.9 where more than one option is available to derive the infinitive of the 2nd conjugation for nonce verb forms *molgo* and *illipingo*.

6 *The learning and generalization of variable morphological rules: the case of the subjunctive*

Four phonotactically illegal verbs were added to the stimuli. They were taken from Say (1998) and used also in Albright (2002a, p. 10): *zuisd-* [dzuizd], *pebcr-* [pebkr], *ghirilb-* [girilb], and *grobm-* [grobm]. Overall, 90 novel verbs were used in the experiment. A list of all the stimuli is provided in Appendix B.

Table 6.8
Descriptive statistics of the experiment stimuli

Descriptives	1 st conj.	2 nd conj.	3 rd conj.	4 th conj.
Mean	0.79	0.27	0.06	0.11
Variance	0.0048	0.093	0.02	0.06
SD	0.069	0.30	0.14	0.24
Min	0.754	0.008	0.002	0.007
Max	0.993	0.981	0.724	0.992
Range	0.23	0.973	0.722	0.985

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Table 6.9

Rules generated by the MGL to derive infinitive forms of novel verbs ‘molgo’ and ‘illipingo’

Rule Environment	1SG present indicative	2 nd conjugation infinitive	Rule reliability based on token frequency	Rule reliability based on type frequency
X[ɔ]	molgo	mɔlere	0.003	0.008
		mɔldʒere	0.41	0.31
		mɔlgere	0.1	0.05
		mɔλλere	0.38	0.41
X[pin]	illipingo	illipingere	0.1	0.05
		illipindʒere	0.99	0.87
		illipinere	0.003	0.008

The stimuli were divided in two groups: half of them, i.e. 45 novel verbs including two phonotactically illegal items were randomly assigned to the first group, the other half to the second group. In experiment presentation 1, the first group of pseudoverbs occurred together infinitive test sentences, while the second group of pseudoverbs occurred with subjunctive test sentences. In experiment presentation 2, the first group of pseudoverbs occurred together with subjunctive test sentences and the second group of pseudoverbs

6 The learning and generalization of variable morphological rules: the case of the subjunctive

occurred with infinitive test sentences. 18 participants received experiment presentation 1, whereas 17 participants received experiment presentation 2. Appendix C illustrates the experimental design and provides examples of how the stimuli are associated with the sentences depending on the experiment presentation and on the participant.

6.5.1.5 Sentence construction and randomization

For the 90 experiment stimuli, 90 test sentences and 100 input sentences were constructed. Half of the test sentences required a 1SG subjunctive present²¹, the other half elicited infinitive forms.

Special care was taken to design test sentences that elicit a subjunctive. As it was discussed in the previous chapters, the subjunctive is often substituted with the indicative in speech. I decided to use only two governors, those that were shown to consistently select the subjunctive in the C-ORAL-ROM corpus (Cresti & Moneglia, 2005): *bisogna che* “it is necessary that” and *vuole che* lit: “(she/he) wants that”. This was done to increase the probability that the participants would produce a subjunctive rather than an indicative or another non-standard form. Half of the subjunctive sentences started with a) while the other half had b):

a) *Bisogna che io* _____

²¹ Unlike other subjunctive forms, the 1sg. present is not syncretic with the indicative counterpart. Therefore, the researcher could establish with absolute certainty whether an indicative or a subjunctive was used. Also, the 1sg. present, unlike other forms, has a different affix for the regular and irregular conjugation classes.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

“It is necessary that I _____”

b) *Vuole che io* _____

“(He/She) wants that I _____”

Both the input sentences and the test sentences were created so as to avoid specific semantic associations with existing verbs. Direct objects were avoided.

Verbs of each group were randomly assigned to the test sentences for each participant. Furthermore, input sentences were randomly assigned to the test sentences for each participant. This was done to further minimize possible semantic associations with existing verbs generated by the sentence context. The presentation of the stimuli was randomized for each participant in order to avoid patterns due to the order of response, fatigue, or level of awareness etc.

10 sentences eliciting past participle forms acted as fillers. They occurred each after 9 test sentences approximately in random order for each participant. 10 novel verbs appeared with the filler sentences. They were randomly selected out of the 90 stimuli for each participant. Eliciting different verb forms, infinitive, subjunctive and past participle, ensured careful appraisal of the test sentences in order to minimize habituation effects.

6.5.1.6 Data analyses

The first objective of the study was to find out whether similarity to existing verbs influences production of novel forms. These measures are captured by the MGL reliability scores.

Thus, the MGL scores were centered and entered as a predictor in a series of generalized linear mixed models (GLMMs) (Bates et al., 2014). Following Veríssimo & Clahsen (2014 p. 68), in the first four regressions the MGL reliability scores for the infinitive rules of each of the four conjugations were entered as simultaneous predictor variables. In this way the independent contribution of phonological similarity to each conjugation could be evaluated while controlling for similarity to other conjugations. Similarity to the competing conjugations may have a negative effect as was the case in Veríssimo and Clahsen's experiment. They argue that the inhibiting effect may arise simply "by virtue of the mutual exclusivity of the different possible responses" (Veríssimo & Clahsen, 2014, p. 68)

The dependent variable in the four regressions was constituted by the infinitive responses of each of the four conjugations taken separately: the first regression predicted 1st conjugation responses (-are), the second regression predicted 2nd conjugation responses (-ere), the third regression predicted 3rd conjugation responses (-ére) and the fourth regression predicted 4th conjugation responses (-ire).

The same procedure was repeated for the subjunctive. The MGL reliability scores for the subjunctive rules of each conjugation (default conjugations and non-default conjugations) were entered as

6 The learning and generalization of variable morphological rules: the case of the subjunctive

simultaneous predictors in the two regressions. The dependent variable in the first regression was constituted by 1st-conjugation subjunctive responses, the verb forms ending in *-i*. In the second regression, the subjunctives of the other conjugations, i.e. all the verb forms ending with suffix *-a* acted as the dependent variable.

The study also aims at uncovering which frequency measure, if any, influences the way in which morphological processes are applied to new words. At first, all three types of reliability scores produced by the MGL were entered as predictors in the regression models: the scores calculated based on the verbs' type frequency, the reliability scores based on the verbs' token frequency and the scores obtained through a combined measure of type and token frequency. However, the three predictors turned out to be highly collinear in the regression²². This means that they explain some of the same variance in the dependent variable²³. For this reason, they cannot be used together in the model

²² The multicollinearity among the predictors was established by calculating the variance inflation factor (VIF) following Marquardt (1970), Belsley et al. (1980) and Zuur et al. (2009, p. 386).

²³ Multicollinearity indicates “when two or more predictors in a regression are highly related to one another, such that they do not provide unique or independent information to the regression” (Taylor, 2020). The example usually provided to explain multicollinearity is using both height and weight as predictors in a regression: “Although they represent different aspects, both are measures of a person's size and one is highly related to the other (taller people are likely to weigh more than shorter people, on average)” (Taylor, 2020). In the model, type frequency and token frequency are related such that, on average, the morphological patterns with high type frequency also exhibit high token frequency.

Detecting multicollinearity is important because it reduces the statistical significance of the independent predictor variables: the estimate of effects for individual predictors are no longer reliable (Lafi & Kaneene, 1992).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

and must be dropped (see Zuur et al., 2009, p. 386). Appendix D shows the correlation coefficients between the predictor variable based on type frequency and the predictor variable based on token frequency for the infinitives (Appendix Table 7) and the subjunctives (Appendix Table 8).

In the regression sets presented in the next section, I dropped the predictor based on token frequency and the predictor based on a combined measure of type and token frequency. The MGL reliability scores entered in the models are based solely on type frequency following Albright (2002a) and Veríssimo & Clahsen (2014). This makes it possible to directly compare the results of the present study with the previous experiments (see section 6.4.1). The statistical analyses were conducted with R Studio (Version 0.99.896, 2009-2016 RStudio, Inc.) (Allaire, 2012) using the `glmer` function of the `lme4` package (Bates et al., 2007; Bates et al., 2012). The models were fitted using step-down regressions starting with the maximal model and progressively eliminating random factors and fixed factors to get the minimum adequate model (MAM). The best model was selected using parametric bootstrap and comparing AIC values as indicators of model fit (following Zuur et al., 2009, see also Wagenmakers & Farrell, 2004 and Vrieze, 2012). Finally, the aim of the experiment was to assess which of the three MGL measures (type frequency, token frequency and type*token frequency) best approximates human responses. To this end, the distribution of infinitive and subjunctive responses for each conjugation given by experiment participants was compared to the

6 The learning and generalization of variable morphological rules: the case of the subjunctive

distribution of infinitives and subjunctives predicted by the MGL algorithm based on the different frequency measures. Comparisons between the predictions of the computational model and human data are operated with the Chi-square statistic²⁴ (Cochran, 1952).

6.6 Results

All the models presented have significant by-participant and by-item random intercepts. No significant random effects of order of presentation of the verbs or of the sentences were detected. Possible semantic effects arising from the sentence were successfully suppressed.

6.6.1 Production of novel infinitives

6.6.1.1 1st-conjugation infinitives

In Table 6.10 the coefficients of the minimum adequate model (MAM) that predict infinitive responses of the 1st conjugation (-are) are presented. According to the model, the probability that the regular infinitive pattern -are occurs increases as the MGL rule reliabilities for this pattern increase. On the other hand, the probability of selecting -are decreases as the MGL rule reliabilities for the competing patterns

²⁴The chi-square test (Pearson, 1989) is used to establish whether there is a statistically significant difference between the expected frequencies and the observed frequencies of a variable. In the present case, we are calculating whether the expected results calculated by the MGL model significantly diverge from the responses given by human participants in the experiment.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

of 2nd and 4th conjugation increase. Thus, even while controlling for the effects of the competing conjugations, a clear effect of similarity to verbs of the 1st conjugation is established. Similarity to the verbs of the 4th conjugation appears only marginally significant. Nevertheless, the best model keeps this predictor.

Table 6.10
Coefficients of the MAM predicting 1st conjugation infinitives (-are)

	Estimated Std.	Error	z value	Pr(> z)
(Intercept)	0.3016	0.2579	1.1169	0.242257
MGL 1 st conj. (-are)	7.0504	2.0431	3.451	0.000559 ***
MGL 2 nd conj. (-ere)	-1.0288	0.4703	-2.187	0.028716 *
MGL 4 th conj (-ire)	-0.8533	0.4953	-1.723	0.084838 .

6.6.1.2 2nd-conjugation infinitives

The best fitting model that predicts infinitives of the 2nd conjugation has the MGL rule reliabilities for the 2nd and 1st conjugation as fixed terms (see Table 6.11). Similarity to 2nd conjugation verbs significantly increases the probability of producing a 2nd conjugation infinitive.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Similarity to 1st conjugation verbs inversely correlates with the probability of *-ere* infinitives.

Table 6.11

Coefficients of the MAM predicting 2nd conjugation infinitives (-ere)

	Estimated Std.	Error	z value	Pr(> z)
(Intercept)	2.7158	1.3291	2.043	0.04101 *
MGL 1 st conj. (-are)	-4.9750	1.5906	-3.1208	0.00176 **
MGL 2 nd conj. (-ere)	1.3637	0.3347	4.074	4.61e-05 ***

6.6.1.3 3rd-conjugation infinitives

Results of the logistic regression predicting infinitives ending in *-ére* with stressed theme vowel show that similarity to verbs of this conjugation increases the probability of *-ére* responses. The best fitting model according to AIC values keeps the MGL reliabilities for other conjugations even if they are non-significant terms (see Table 6.12). It must be borne in mind that very few verbs belong to this conjugation overall (only 1.1%) and participants very rarely gave *-ére* responses. Therefore, the uncertainty in the sample and in the estimation is great.

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Table 6.12
Coefficients of the MAM predicting 3rd conjugation infinitives (-ére)

	Estimated Std.	Error	z value	Pr(> z)
(Intercept)	-0.3473	15.4166	-0.022	0.9820
MGL 1 st				
conj. (-are)	-8.5409	18.4808	-0.462	0.6440
MGL 2 nd				
conj. (-ere)	-2.5644	4.9034	-0.523	0.6010
MGL 3 rd				
conj. (-ére)	6.8064	2.8950	2.351	0.0187 *
MGL 4 th				
conj. (-ire)	-40.1923	169.2821	-0.237	0.8123

6.6.1.4 4th-conjugation infinitives

Finally, the best fitting model for the infinitive responses of the 4th conjugation (-ire) confirms that the similarity of nonce verbs to lexical items of this class increases the probability of producing -ire infinitives (see Table 6.13).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Table 6.13

Coefficients of the MAM predicting 4th conjugation infinitives (-ire)

	Estimated Std.	Error	z value	Pr(> z)
(Intercept)	-3.2551	0.3217	-10.120	< 2e-16 ***
MGL 2 nd conj. (-ere)	-1.0580	0.5829	-1.815	0.069529 .
MGL 4 th conj. (-ire)	1.9219	0.5152	3.731	0.000191 ***

6.6.2 Production of novel subjunctives

For the test sentences that elicit subjunctive forms, only two governors were used, *bisogna che* “it is necessary that” and *vuole che* “he/she wants that”. These were the governors that showed the least amount of mood variation in the corpus study and consistently selected the subjunctive. The substitution of the subjunctive with other moods was successfully minimized in the experiment. Only one male participant (aged 51) of the 35 adult native speakers produced exclusively indicative forms in subjunctive sentences. These forms were excluded from the analysis. The remaining 1573 subjunctive observations were entered as a dependent variable into two generalized linear mixed models.

6.6.2.1 1st-conjugation subjunctives

The results of the logistic regression (Table 6.14) confirm a similar pattern to the one obtained for the infinitives. Similarity to existing verbs of the 1st conjugation greatly increases the probability of producing a 1st-conjugation subjunctive while similarity to the other classes inhibits 1st-conjugation responses, an effect that is only marginally significant.

Table 6.14
Coefficients of the MAM predicting 1st conjugation subjunctive (-i)

	Estimated Std.	Error	z value	Pr(> z)
(Intercept)	-6.2757	1.8972	-3.308	0.000940 ***
MGL 1 st conj. (-i)	7.3826	2.1592	3.419	0.000628 ***
MGL non-default conj. (-a)	0.9409	0.5054	-1.862	0.062619 .

6.6.2.2 Subjunctives of the non-default conjugation classes

According to the minimum adequate model fitted to the subjunctive forms of the non-default classes, phonological similarity to verbs of the 1st conjugation reduces the probability of producing non-default

6 The learning and generalization of variable morphological rules: the case of the subjunctive

subjunctives. Similarity to verbs of the non-default classes positively influences this probability, an effect that is only marginally significant (see Table 6.15).

Table 6.15
Coefficients of the MAM predicting subjunctives of the non-default classes, 2nd, 3rd and 4th conjugations (suffix -a)

	Estimated Std.	Error	z value	Pr(> z)
(Intercept)	6.3988	1.9163	3.339	0.000841 ***
MGL 1 st conj. (-i)	-7.8556	2.1804	-3.603	0.000315 ***
MGL non-default conjugations (-a)	0.9513	0.5081	1.872	0.061147 .

6.6.3 The role of type and token frequency in morphological productivity

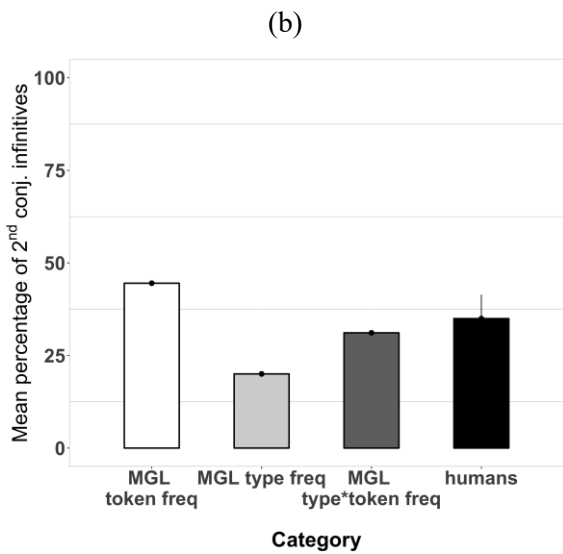
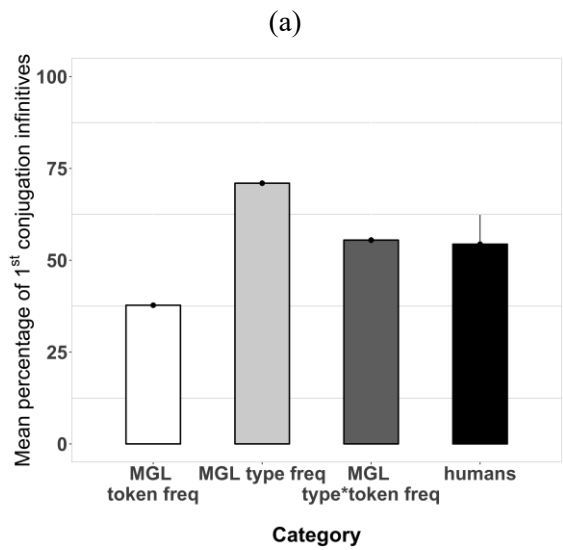
This section addresses the predictions of the network model and of connectionist models concerning which frequency measure influences the productivity of morphological patterns. The distribution of infinitive and subjunctive responses for each conjugation was compared to the distribution estimated by the MGL algorithm based on the different frequency measures. The aim is to find out which of the MGL measures best approximates human behavior. Comparisons are operated with the Chi-squared statistic (Cochran, 1952).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

Recall that type frequency reflects the number of different words with which a given morphological pattern occurs whereas token frequency reflects the absolute frequency with which a morphological pattern occurs.

A combined measure of type and token frequency best approximates human responses for the 1st and 2nd conjugation infinitives (see Figure 6.1 (a) and (b)). This is the only measure for which the Chi-squared statistic does not report any significant differences between human responses and the predictions of the computational model at a $p < .01$ level of significance: $\chi^2(1, N = 1,576) = .28, p = .59$ for 1st conjugation infinitives and $\chi^2(1, N = 1,576) = .485, p = .027$ for 2nd conjugation infinitives. Type frequency overestimates the application of the 1st conjugation pattern to new words while token frequency underestimates it. The opposite happens with 2nd conjugation infinitives: type frequency alone greatly underestimates the generalization rate of this pattern by experiment participants while token frequency overestimates it. For the 3rd and 4th conjugation infinitives, type frequency alone appears to best capture the responses given by native speakers, $\chi^2(1, N = 1,576) = 1.03, p = .308$ for 3rd conjugation infinitives and $\chi^2(1, N = 1,576) = .07, p = .077$ for 4th conjugation infinitives. However, these conjugations are so small that very few responses were given overall. The uncertainty in the sample and that of the model estimates is much greater (see Figure 6.1 (c) and (d)). Therefore, the outcomes are less reliable.

6 The learning and generalization of variable morphological rules: the case of the subjunctive



6 The learning and generalization of variable morphological rules: the case of the subjunctive

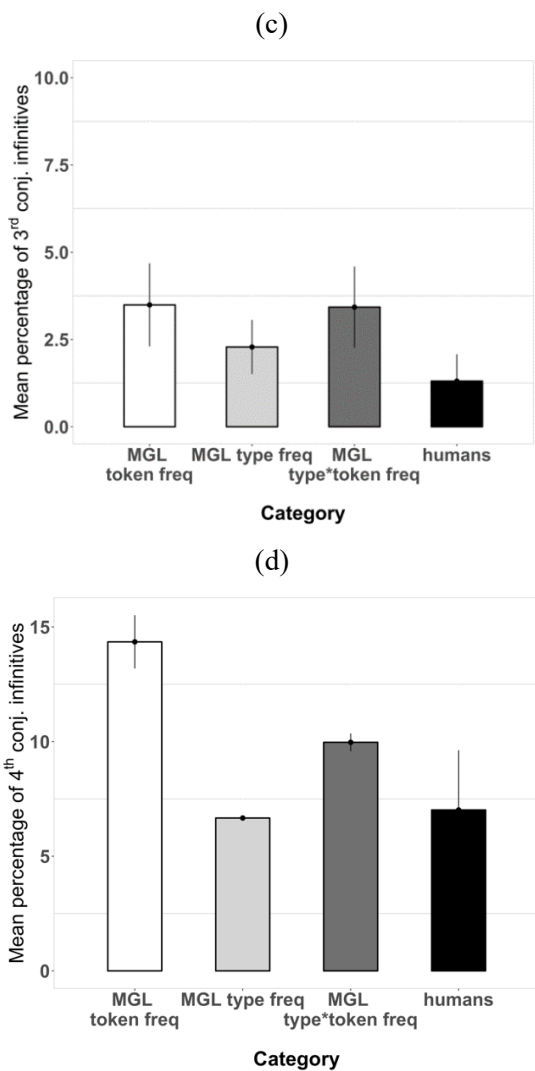


Figure 6.1: Mean percentages of infinitive responses with associated 95% confidence intervals obtained with human participants and predicted by the MGL based on various frequency measures (token frequency, type frequency and type*token frequency).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

The results for the subjunctives confirm what was observed for the infinitives: the estimates of the MGL model by means of a combined measure of type and token frequency best predict human performance (see Figure 6.2). They are statistically equivalent to the rates produced by the experiment participants, $\chi^2(1, N = 1,576) = .09, p = .763$ for 1st conjugation subjunctives and $\chi^2(1, N = 1,576) = 2.1, p = .147$ for the subjunctive forms of the non-default conjugations. Crucially, in the case of the subjunctive, type frequency alone greatly overestimates the productivity rates of 1st conjugation responses (Figure 6.2 (a)) and greatly underrates the productivity of 2nd, 3rd, and 4th conjugation responses (Figure 6.2 (b)).

Token frequency appears to be even more relevant as a driver of morphological productivity. Indeed, a comparison between the productivity rates of the subjunctive and infinitive patterns for each conjugation (shown in Figure 6.3) reveals a striking fact. The infinitive responses match what has been written about the conjugations' productivity and default status. The 1st conjugation is the only productive and default verb class nowadays. It is no surprise that the great majority of infinitive responses exhibit the 1st conjugation pattern. The infinitive responses that exhibit the morphological patterns of the other conjugation classes are far fewer. Yet with the subjunctives the productivity rates shift dramatically. Subjunctive morphology belonging to the conjugation class with lower type frequency becomes the most productive: suffix *-a* is used with novel verbs to derive the subjunctive significantly more often than suffix *-i*. The very same

6 The learning and generalization of variable morphological rules: the case of the subjunctive

speakers who almost exclusively produced 1st conjugation infinitives, almost exclusively chose the non-default pattern for the subjunctive.

This is the morphological pattern that appears with a small number of high-frequency irregular verbs, which, taken together, constitute 66.8% of the subjunctive occurrences in the corpus (see again Table 6.6).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

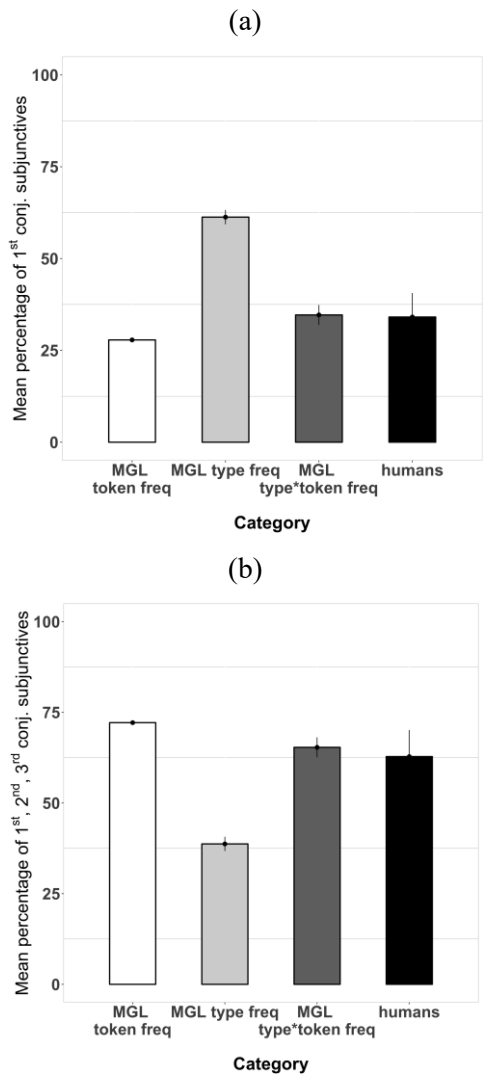


Figure 6.2: Mean percentages of subjunctive responses with associated 95% confidence intervals obtained with human participants and predicted by the MGL based on various frequency measures (token frequency, type frequency and type*token frequency).

6 The learning and generalization of variable morphological rules: the case of the subjunctive

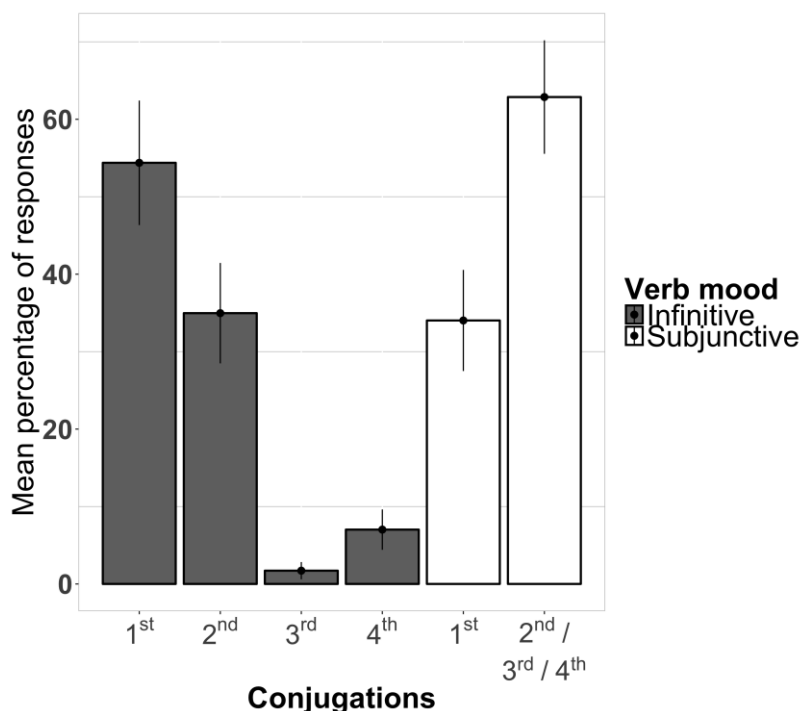


Figure 6.3: Comparison between infinitive and subjunctive productivity rates for each conjugation.

6.7 Discussion

The experiment addressed the hypotheses of the dual-mechanism model and those of the single-route analogy-based models concerning the mechanisms in which morphological rules are applied to new words. It was established that adult native speakers are sensitive to the phonological shape of words when selecting a rule among different options. In particular, the application of morphological patterns to new words is driven by phonological similarity to existing words and

6 The learning and generalization of variable morphological rules: the case of the subjunctive

sensitive to the statistics of the lexicon. This is the case for both the regular productive pattern, i.e. infinitives and subjunctives of the 1st conjugation, and irregular or unproductive ones (subjunctives and infinitives of the 2nd, 3rd and 4th conjugations).

The results favor single-route analogy-based models and are in line with evidence gathered by Albright (2002a) and Albright (2002b) on the acceptability rates of novel infinitives by Italian native speakers. They directly contradict the predictions of the dual-mechanism model, which assumed that the application of regular 1st conjugation morphology is insensitive to the phonological properties and the statistics of existing verbs and takes place only by means of a context-free rule of the type *Y-are* (see Say & Clahsen, 2002 and Veríssimo & Clahsen, 2014). Also, the outcome of the experiment does not conform with evidence in Veríssimo & Clahsen (2014) who report a dissociation between the generalization of the 1st conjugation and that of the other conjugations in the production of novel infinitives in Portuguese. This is somewhat puzzling as the design of the latter experiment is very similar to the one adopted in the present study. In section 6.4.1 I discussed one striking difference in the reliability score of the context-free rule discovered by Veríssimo and Clahsen for the regular infinitive pattern of the 1st conjugation in Portuguese (rule *Y-ar*, reliability score smaller than or equal to .459) and the one reported in Albright (2002a) and myself for Italian (rule *Y-are*, reliability = .747). This difference of .288 points (or more) does not seem to reflect the size of the conjugation classes which are very similar in Italian and Portuguese. Indeed, the 1st

6 The learning and generalization of variable morphological rules: the case of the subjunctive

conjugation class appears to be even larger in Portuguese (82.3% of Portuguese verbs fall in the 1st class vs. 74.9% Italian verbs, again see Table 6.3). The discrepancy must be addressed in future research in order to compare the verb morphology of the two languages directly.

Additionally, Veríssimo and Clahsen do not report the use of filler sentences in their experiment. It appears that only infinitives were elicited. This may have triggered habituation effects whereby participants no longer engage in the task with a high level of attention and tend to give the same response as a result. This constitutes another major difference between their experiment and the present task, in which test sentences alternated regularly between infinitive, subjunctive and past participle contexts.

Having established that the statistical properties of the lexicon influence the extension of morphological patterns to new words, the study aimed at clarifying a long-standing debate between analogy-based models on which type of statistics directly influences the productivity of morphological patterns.

The experiment reveals that **both type and token frequency** are crucial to the extension of morphological patterns to new words in this type of task. This result supports connectionist models (Rumelhart & McClelland, 1986) and contradicts the hypothesis of the network model (Bybee, 1995) according to which only type frequency is relevant for the productivity of a pattern. Token frequency is even assumed to detract from its productivity in this model (Bybee, 1995, p. 435). The network model is largely based on the morphology of English where

6 The learning and generalization of variable morphological rules: the case of the subjunctive

productivity and irregularity coincide (Baayen, 1992, p. 140). The case of the subjunctive however clearly shows how important token frequency is: the subjunctive morpheme belonging to the conjugation classes with lower type frequency becomes the most productive in the experiment. This is the morpheme that appears in the subjunctive form of **high-frequency, irregular verbs**. They all take segmentable affix -*a*:

Infinitive		Subjunctive 1 st /3 rd sg. present tense
<i>venire</i> ‘to come’	→	<i>venga</i>
<i>andare</i> ‘to go’	→	<i>vada</i>
<i>essere</i> ‘to be’	→	<i>sia</i>
<i>avere</i> ‘to have’	→	<i>abbia</i>
<i>potere</i> ‘can’	→	<i>possa</i>
<i>dire</i> ‘to say’	→	<i>dica</i>

A possible criticism against this type of experiment is that they do not really capture how morphological productivity works in natural languages. Adult speakers may surmise the goal of the task and strategize in the experimental setting. Children on the other hand are more naïve with regards to experimental manipulation when they perform wug experiments. Their linguistic behavior is assumed to reflect more realistically how morphological rules are learned and generalized (Pinker, 1995a).

Based on evidence coming from wug experiments (Berko, 1958; Xu & Pinker, 1995b), Yang (2016) notes that the behavior of children differs

6 The learning and generalization of variable morphological rules: the case of the subjunctive

from that of the adults: “[c]hildren generalize productive rules but do not generalize lexical rules” (p. 1168).

An interesting venue for future research would be to repeat the experiment with children and establish to what extent their linguistic behavior differs from that reported for adult speakers. Through the experiment it was established at the very least that adult speakers store gradient and fine-grained statistical and phonological information concerning the linguistic input they receive and are indeed able to use it in the context of deriving morphologically complex words.

7 Conclusion and general discussion

7.1 Summary of the results

This dissertation has analyzed the perception, usage and productivity of the Italian subjunctive. Due to its peculiar characteristics, the subjunctive was considered an interesting case study on which to investigate how productive morphological rules are formed and maintained in the face of morphological variation, simplification and obsolescence. The present work integrates very different strands of linguistic research, reviewing and critically evaluating conflicting views in order to give a comprehensive analysis of the subjunctive that has not yet been provided.

In the introduction, I presented the notion of mood variation. Mood variation indicates the alternation between the subjunctive and other mood forms, such as the indicative or the conditional in the utterances that would select subjunctive morphology according to the prescriptive language norm. It is mainly found in speech although it is also common in informal written communication.

In chapter 2 it was shown that the Italian linguistic community at large perceives the substitution of the subjunctive with other moods as an error and it is one that generates a lot of attention, strong feelings and opinions. Indeed, it was discussed how the consciously perceived aspects of this phenomenon create metalinguistic discourses embedded in songs, books, films, slogans and online discussions etc. These discussions have a variety of purposes, like sanctioning non-standard

7 Conclusion and general discussion

language practices, being ironic, generating stereotyped identities, and making moral comments about socio-cultural behaviors, single individuals or social groups. For example, at different points they have been used to parody the lower class, the middle class, the political class and even men in general etc. The social phenomena that accompany synchronic morphological or phonological variability are extensively studied in sociolinguistics and folk linguistics (e.g. Silverstein, 2003; Bucholtz & Hall, 2004; Niedzielski & Preston, 2010). Another type of error that receives equal attention from non-linguists is the production of so-called “distorted subjunctives”: verb stems of the 1st conjugation which are combined with subjunctive affixes belonging to other conjugations and vice-versa. The outcome are forms that do not exist in Italian. Examples are **venghi* for *venga* “come”, **vadi* for *vada* (1/3SG SUBJ PRS), *venghino* for *vengano* (3PL SUBJ PRS), *trovessimo* for *trovassimo* “find” (3PL IMPF SUBJ). Only the overgeneralization of the morphemes of the 1st conjugation is reported in the literature. This phenomenon is taken to be the result of morphological simplification within the subjunctive paradigm which involves analogical leveling to the forms of the 1st class, the productive conjugation that groups most verbs (see Treccani, 2012).

Only certain aspects of this morphological variability receive the attention of non-linguists. For instance the use of conditional forms in lieu of imperfect subjunctives in conditional sentences is disparaged and parodied. Other non-standard practices are largely uncommented indicating that they have been normalized.

7 Conclusion and general discussion

Although the public ear is sensitive to these grammatical errors, it was shown that they are pervasive in speech. Indeed, they also occur in very formal speech and in partially scripted talk, for instance in speeches given by politicians in official settings. Such phenomena raised questions concerning the speakers' internal grammar of the subjunctive:

1. How can the use of the subjunctive be described and understood?
2. Is the morphological variation presented completely random?
 - a. Do semantic factors accurately describe the subjunctive/indicative distribution in speech?
 - b. Is the subjunctive fossilized in a number of subjunctive phrases or *schemas*?

Additionally, the distorted subjunctive produced by native speakers point to issues of productivity of the subjunctive morphemes. The appropriate morpheme simply does not occur to speakers on the fly, while articulating. The study addresses the following questions:

3. Is the subjunctive overall still productive in Italian?
4. Are some forms of the mood system disappearing faster than others?

Ultimately, the subjunctive was used as a case study in order to explore how frequency impacts morphological learning and productivity.

7 Conclusion and general discussion

Chapter 3 discusses the extensive studies of the subjunctive mood conducted within the framework of descriptive linguistics, semantics, modal logic, generativist theory and the minimalist program among others (e.g. Picallo, 1984; Farkas, 1992a, 1992b; Quer, 1997, 2009; Giorgi and Pianesi, 1997, 2002, 2004a, 2004b; Giorgi, 2009; Kempchinsky, 1987). This line of work starts from the assumption that mood distribution is semantically motivated. It aims to find the core semantics of the subjunctive and formulate very general rules that explain its peculiar morphosyntactic distribution. In more traditional accounts, the subjunctive is presented as encoding *irrealis*, non-declarative, non-factual or subjective utterances in binary opposition with the indicative, the mood of the *realis*, declarative, factual and objective propositions. The same modal meanings are assumed to be encoded in the subjunctive that occurs in subordinate clauses. Indeed, the modal value of the subjunctive must be consistent with the meaning of the governors that select it (Farkas, 1992b). Some scholars argue that it is impossible to treat the subjunctive as a unitary notion and they develop different semantic categories that are meant to capture the peculiar morphosyntactic behavior of the subjunctive in the various categories (see Quer, 1997, 2009; Prandi, 2002). Oftentimes, semantic categories are derived from the classification employed for the subjunctive in Latin. For instance, we have categories like:

i. *Polarity subjunctive*

It is meant to indicate contexts where the negation of the matrix affects the indicative/subjunctive distribution, mainly with

7 Conclusion and general discussion

epistemic verbs like *believe*, *think*, *seem* etc. In such cases, negation is taken to affect the epistemic commitment to the truth of the proposition. It conveys non-realistic, improbable eventualities encoded by the subjunctive.

ii. *Epistemic subjunctive*

It is governed by verbs that express a weaker epistemic commitment with regard to the proposition in the complement clause (Farkas, 1992b, p. 71), verbs like *believe*, *imagine*, *think*, *admit*, *suppose*, which convey hypotheses, doubts and beliefs, etc.

iii. *Optative or volitive subjunctive*

It is licensed by desiderative verbs encoding a want or a wish that the event of the embedded proposition comes to pass (Renzi et al., 1991, p. 416).

Finally, work within the generativist tradition strived to capture the morphosyntax of the subjunctive under a unitary theory. In particular, it addressed:

- a) The temporal interpretation of subjunctive clauses
- b) The constraints on co-reference of subjunctive clauses
- c) The role of the complementizer in subjunctive clauses

In regard to a), the *consecutio temporum* constraint was explored. According to this constraint, the tense of the embedded subjunctive must agree with the tense of the matrix clause. It was claimed that the

7 Conclusion and general discussion

subjunctive, unlike the indicative, is inherently tenseless (Picallo, 1984). It can refer to the tense of the matrix clause but it cannot independently encode temporal information (Giorgi, 2009). In this respect the subjunctive is understood to have properties of non-finite moods.

Concerning b), certain subjunctive clauses do not allow coreference between the matrix subject and the subject of the embedded clause. This is known as disjoint reference constraint (or disjoint reference effect) in the literature (Bianchi, 2001; Kempchinsky, 2009).

With regard to point c), a well-studied syntactic property associated with the subjunctive is the possibility to omit the complementizer “that”, which connects the main clause to the embedded argument clause. Complementizer deletion (CD) is not possible when the embedded clause selects the indicative (Giorgi & Pianesi, 1997, 2004; Poletto, 2000, 2001).

Scholars noticed that all these syntactic phenomena correlate. Usually, when *consecutio temporum* is enforced, the subjunctive disjoint reference effect is also active and the complementizer can be deleted. Per contra, when the subjunctive clause violates the *consecutio temporum*, the subjunctive disjoint reference can be violated and the complementizer cannot be deleted.

Largely simplifying, all these phenomena are explained by the different theories drawing on the notion of a viewpoint or perspective that the sentence encodes. The viewpoint is “the source of the report, the person with respect to whose consciousness (or “self”) [...] and [...]

7 Conclusion and general discussion

from whose point of view the report is made” (Sells, 1987, p. 445). The subjunctive encodes the viewpoint of an internal subject, an antecedent from whose mental state the utterance is evaluated. This is distinct from the viewpoint of the external speaker or hearer encoded in the indicative. For this reason, the subjunctive has been called a *logophoric mood* (Bianchi, 2001, p. 36). This internal perspective is also claimed to be encoded in non-finite clauses.

Throughout chapter 3, ample evidence coming from language corpora was provided to show that such theories fail to capture the mood distribution in speech. Oftentimes the indicative appears in semantic contexts, in which the afore-mentioned theories predict the occurrence of a subjunctive. Moreover, it was shown how these very abstract and complex theories are actually inconsistent with each other even when they draw on the very same concepts (see Table 3.2).

Against this backdrop, more data-driven research, including my own analyses, shows how the alternation between the subjunctive and other moods, like the indicative and the conditional, cannot be properly grasped by relying exclusively on semantic-based accounts. They provide evidence that the subjunctive in subordinate clauses functions as a mere syntactic maker of subordination, devoid of any meaning relating to modality (Harris, 1974; Bergareche, 1990; Poplack, 1990; Silva-Corvalán, 1994a, 1994b; Bybee 2002b). Some of these studies have proposed that, in the diachronic development from Latin to the Romance languages, the subjunctive has progressively lost semantic and grammatical meaning, a process known as desemanticization and

7 Conclusion and general discussion

viewed as part of grammaticalization (see section 5.1.1). The implicit assumption is that the subjunctive in Latin had a core modal meaning (see Poplack et al., 2018).

Chapter 4 addressed this hypothesis by reviewing the studies on the functions and syntax of the subjunctive in Latin and comparing them to the use of the mood in the Romance languages. In this way, indirect clues to the development of the subjunctive were gathered. Additionally, the metalinguistic reflection of the ancient Roman writers and grammarians concerning the subjunctive in classical Latin was reviewed and compared to the theoretical work discussed in chapter 3 in order to supply a relevant context for the theories that attempted to explain the mood opposition in more recent times. Ancient grammarians already presented the subjunctive as the mood of the hypotaxis and reported that it was dependent on additional lexical elements like the subordinating conjunction and the governor verb to acquire meaning (see the excerpts from Macrobius, Cledonius and Priscian among others in section 4.1). Thus, finding the core underlying grammar of the subjunctive that may account for all of its surface realizations seems to be a preoccupation of more recent times probably spurred by theoretical frameworks such as generative grammar. The subjunctive conveyed different functions in unembedded clauses in Latin. If they are kept in the Romance languages, they are largely fossilized. In most of them, it was replaced by other mood forms, like the imperative, the conditional and the infinitive, or by lexical periphrases which became more productive. Oftentimes the mood

7 Conclusion and general discussion

becomes relegated to the subordinate clause which takes up the functions that used to be encoded by the unembedded subjunctive (see for instance the Latin *concessive subjunctive* which gave way to the concessive subordinate clause where the subjunctive is licensed by the subordinating conjunction).

Thus, it appears from a mere qualitative analysis that in Latin the subjunctive did indeed have a number of independent meanings and was more productive in unembedded clauses than in its daughter language today. However, a quantitative investigation of its distribution shows that its use in unembedded utterances was on the whole relatively rare. It was far more frequent in subordinate clauses. In these contexts, its functions were so varied, ranging from the semantic to the fully grammatical, that it would have been impossible for speakers of the time to gather a core meaning (Pinkster, 2015, p. 617). Most importantly, in many cases it already acted as a mere sign of subordination and was lexically selected by a governing element (see for instance the subjunctive in indirect questions). Still, in Classical Latin the subjunctive was the most frequent mood after the indicative. This suggests an overall extremely large decrease in frequency in relation to other moods in the Romance languages. Mood forms like the infinitive, the conditional or the imperative gradually increased in frequency and took up some of the functional domains previously conveyed by the subjunctive.

Having established the historical background, chapter 5 takes on the task of mapping mood variability in Italian and of identifying the

7 Conclusion and general discussion

factors that prompt the selection of the subjunctive in speech. A corpus study was conducted using the Italian C-ORAL-ROM (Cresti et al., 2002). The goal of the research was to explore the influence of the governor on mood distribution in subordinate clauses. The statistical analysis revealed that no single governor exclusively selects the subjunctive, not even the desiderative/volitive verbs, which, according to some accounts, are considered to represent the core subjunctive meaning (see Kempchinsky, 2009). Most interestingly, governors belonging to the same semantic class or quasi-synonymous verbs exhibit different correlations with the subjunctive (see for instance the epistemic verbs). This result defeats theories claiming that mood distribution is mainly semantically motivated and parallels evidence from other corpus studies in other Romance languages (see Poplack, 1991; Poplack et al., 2013; Silva-Corvalán, 1994b; Bybee & Thompson, 1997; Montrul, 2007 among others). Moreover, only two governors consistently select the subjunctive nowadays. They are the verbs *volere* “to want” and *bisognare* “to necessitate” in the form *bisogna* “it is necessary”, the same two that strongly correlate with the mood in French (see Poplack et al., 2013). This evidence suggests that the subjunctive occurs in formulaic phrases or “schemas” (see Bybee & Thompson, 1997), a result of a process known as obligatorification in historical linguistics or as entrenchment in the cognitive literature: lexical items that frequently occur together become routinized and are perceived and processed as a tighter unit. Obligatorification is considered part of grammaticalization and an indicator of a decrease in

7 Conclusion and general discussion

productivity. Crystallized phrases ensure that the obsolescing grammatical form is still produced. However, speakers no longer use it freely to create new meaning.

Concerning the role of verb frequency in the mood distribution, a comparison with the infinitive revealed that the subjunctive disproportionately occurs with a small number of high-frequency verbs. Only five verbs represent 66.8% of the subjunctives vs. only 29.6% of infinitives in the C-ORAL-ROM corpus of spoken Italian (see Table 5.2). This particular distribution was also found in French (Poplack et al., 2013) and gives further evidence that obsolescing or unproductive morphological patterns are better preserved in high-frequency lexical items (see Bybee & Slobin [1982] and Lieberman et al. [2007] among others).

Chapter 6 investigates the productivity of the subjunctive paradigm of each conjugation with the use of a wug experiment. The goal of the study was to address long-standing research questions on the mental representation of morphological knowledge and on the factors that influence morphological productivity.

In particular, the following questions were explored:

- d) How do speakers generalize morphological patterns to new words?
- e) Is the generalization of both regular and irregular morphological patterns influenced by similarity to existing words?

7 Conclusion and general discussion

- f) Does frequency of occurrence of the morphological patterns influence their productivity?

Models of morphological learning make different predictions. The dual-mechanism model (Pinker & Prince, 1988; Pinker, 1991) posits that only irregular rules are generalized by means of analogy to existing words and are thus sensitive to the phonological properties and the statistics of the lexicon. Regular morphological processes are applied by means of a context-independent rule that is insensitive to the properties of the lexicon. Analogy-based models, on the other hand, assume that the experience with language that the child and the adult speaker receives is fundamental in shaping her morphological knowledge (Rumelhart & McClelland, 1986, McClelland & Patterson, 2002; Bybee, 1995). In particular, all morphological rules, productive and unproductive, regular and irregular, are formed by extracting similarities between existing words. The frequency with which a given morphological change occurs in the input is fundamental in determining whether the rule will be productive and will hence be extended to new words. Analogy-based models diverge on the type of frequency that directly influences productivity. According to the network model, only the number of different words that take a morphological pattern directly influences its productivity. This measure is captured by type frequency. The hypothesis is based on the observation that high-frequency words do not seem to improve the productivity of irregular patterns (e.g. irregular past tense forms in English). Only if the morphological

7 Conclusion and general discussion

schema is represented in a great set of words will it be generalized. Connectionist models on the other hand predict that the absolute frequency with which a morphological change is experienced will influence the probability that it will be applied to derive new word forms. This means that both the number of different lexical items it occurs with (type frequency) and its absolute frequency (token frequency) play a role in its productivity.

These hypotheses were tested by using an oral production task in which speakers were asked to derive subjunctives and infinitives of new words. The two moods were chosen because they differ in crucial aspects concerning the frequency with which they are experienced in language. The majority of subjunctives occur with high-frequency verbs which take the non-default morphological pattern (2nd, 3rd and 4th conjugation). Thus, based solely on token frequency, the non-default pattern would be the most frequent in speech. The default subjunctive pattern (1st conjugation) has the highest type frequency. Per contra, the infinitive pattern of the 1st conjugation has the highest type and token frequency (see Table 6.5 and Table 6.6).

The results of the experiment confirm the predictions of analogy-based models. The application of all morphological processes was found to be influenced by the novel verbs' phonological similarity to existing words. This was the case for both the subjunctives and the infinitives. There was no difference in the way regular and irregular processes were applied, an outcome that undermines the dual-mechanism model.

7 Conclusion and general discussion

Additionally, a **combined measure of type and token frequency** was found to predict the extension of morphological patterns to new words in this type of task. Type frequency alone overestimated the production of the regular 1st-conjugation infinitives and subjunctives. In the case of the subjunctive, the overestimation was sizeable: type frequency was by far the worst measure and performed much worse than token frequency (see Figure 6.2 (a)). The same occurred for the application of the non-default irregular patterns: type frequency underestimated the production of these forms, even more so in the case of the subjunctive (see Figure 6.2 (b)).

This result corroborates the predictions of the connectionist models (Rumelhart & McClelland, 1986) and contradicts the hypothesis of the network model (Bybee, 1995).

Finally, the experiment measured the productivity rates of the subjunctive and the infinitive patterns of the different conjugations. The rates of infinitives of each conjugation produced in the experiment neatly match the sizes of the conjugation classes. The most frequent responses were 1st-conjugation infinitives (-are), followed by 2nd-conjugation infinitives (-ere), 4th-conjugation infinitives (-ire) and lastly by 3rd-conjugation infinitives (-ére) which were extremely few overall. This distribution is completely subverted when considering the subjunctive responses. Strikingly, most subjunctives produced bear the pattern of the smallest conjugations (suffix -a), those that are no longer considered productive in Italian, while the subjunctives of the 1st conjugations (suffix -i) are far less frequent (see Figure 6.3). The non-

7 Conclusion and general discussion

default subjunctive pattern emerges as the most productive in the experiment. The same participants who produced the infinitives of the 1st conjugation shifted to the morphology of the other conjugation classes when producing novel subjunctives. This result would be extremely puzzling if we did not consider that the majority of subjunctive forms occurs with a small number of high-frequency verbs in speech. These verbs exhibit the non-default pattern, i.e. suffix *-a* along with possible stem changes. Thus, Italian native speakers when exposed to the subjunctive, most frequently hear the non-default pattern, even if it is for the most part repeated in only a few lexical items. As it was reported, a combined measure of type and token frequency accurately predicts these rates.

7.2 General discussion and outlook

The dissertation establishes a connection between the patterns that emerge in the corpus study (chapter 5) and the behavioral data gathered in the wug experiment and presented in chapter 6. From the evidence collected, it is possible to conclude that experience with language is crucial in deriving morphological knowledge. More specifically, the linguistic behavior of experiment participants suggests that morphological knowledge is stored in the form of gradient, largely redundant stochastic rules (see Albright, 2002a): they specify, along with the morphological change, information about the context of application of the rule (e.g. verbs that share a particular phonological sequence) and the probability of its occurrence in a given set. This

7 Conclusion and general discussion

information is derived from the distributional properties of complex words and base words in the linguistic input that speakers receive throughout their lifetime. According to connectionist models, an analogical mechanism allows to extract this information (Rumelhart & McClelland, 1986; Eddington, 2002). Analogy is understood as a structure-sensitive mechanism that enables speakers to derive relations of similarity and difference by comparing lexical items together (Gentner & Markman, 1998, 2002).

A context-free rule is a rule that applies to all words of a given syntactic category. For instance, suffix *-are* is added to any type of verb stem to derive the infinitive form in Italian. It is conceptualized as an operation over a variable: Y-are. It is discovered when the morphological change is experienced in so many different words that no sub-regularities can be found by the learner. In the model MGL developed by Albright & Hayes (2003) and in analogy-based models, the context-free rule is also assigned a probability of occurrence and is no different in nature than the context-dependent rules. Productivity is thus a gradient notion. This represents a major contrast between analogy-based models and the dual-mechanism model (Pinker, 1999). In the dual-mechanism model the context-free rule is understood as a symbolic operation and its representation is distinct from that of lexical items (Clahsen, 2006). This rule does not have a probability of application based on the lexicon and for this reason, according to a number of scholars, it is overgeneralized by children in the process of language learning (Marcus et al., 1992; Pinker, 1995b; Yang, 2002).

7 Conclusion and general discussion

Only irregular patterns can sometimes be generalized by means of analogy to existing forms.

Additional studies are needed to evaluate new proposals of how morphological patterns become productive and compare them with established theories. For instance, a very recent and influential account is the Tolerance Principle put forward by Yang (2016).

The Tolerance Principle is based on the relationship between the rule-following items and the exceptions: a productive context-free rule is formed only if the number of exceptions does not surpass a tolerable limit. The Principle ultimately hinges on the statistics of the lexicon. However, once a productive rule is formed, speakers will exclusively use this rule to derive new words. If there are too many exceptional forms, no productive rule can be extracted. Only then will speakers adopt another strategy and find sub-regularities that divide the lexicon into subsets based on their features. This notion is not gradient. There is a specific threshold of productivity. The actual threshold changes as a function of the size of the lexicon and is specified in the following formula (see Yang, 2016, p. 64):

(i)

$$e \leq \frac{N}{\ln N}$$

where e is the number of exceptions and N is the number of words. The number of exceptions must remain below the number of words divided by the logarithm of the number of words.

In the event that the productivity threshold is not reached, more specific context-dependent rules will be formed. They are nested in the

7 Conclusion and general discussion

most general rule. Provided that a given word meets the criteria of the more specific rule, then the latter will automatically apply to it. If not, the more general rule will be applied. Yang (2016) gives the examples of English suffixes used to derive nouns out of adjectives:

[suffix] -ness is the most general rule that has no obvious restriction on the adjective stem, while -ity is nested within, because it productively attaches to adjectives that end in -ible, -ic, and -al. As such, adjectives with these specific phonological properties are handled by -ity and do not constitute exceptions to the -ness suffix. In computational terms, a word like *credible* is shipped to the -ity “subroutine” without ever coming across the -ness rule. These subregularities can only be revealed when learners fail to detect a productive rule over the entire set of adjectives [...] (pp. 73-74).

In the experiment, adult speakers appear to adopt rules in a gradient manner, using both the more general context-free rule and the more specific rules. This result could in theory be compatible with the Tolerance Principle if no productive rule is discovered for the Italian infinitives and subjunctives. This is indeed the case: applying the mathematical formula in (i) to the infinitives and subjunctives reveals that there are too many exceptions to the candidate rule (the 1st-

7 Conclusion and general discussion

conjugation pattern)²⁵. Therefore, more specific rules can be generated such as the ones produced by the MGL algorithm and used to model the behavior of adult participants in the wug experiment. In that case, however, the theory developed by Yang would predict that the more specific rule will always be preferred over the more general. This is not the case for the MGL. The rules are always ordered according to their reliability, i.e. the probability that they will occur in a given set. A more general rule can prevail over more specific ones if it has a higher reliability. The model actually disfavors rules that apply to fewer items (see again section 6.4).

The MGL successfully predicted the subjunctive and infinitive forms produced by adult speakers in the experiment. However, the experiment was not designed to test and compare the specific predictions of the Tolerance Principle against those of the MGL. Additional studies are needed to reveal in more detail how analogical mechanisms use the distributional properties of the lexicon to extract morphological patterns. This can only be done by developing experiments that directly test the different computational models and are able to directly falsify the models examined.

Amenta & Crepaldi (2012) aptly notice that “often new models are put forward without previous models being clearly falsified and without an explicit comparison that could clarify whether and how the new

²⁵ Yang’s (2016) Tolerance Principle was applied to the number of infinitives and subjunctive found in the C-ORAL-ROM corpus (the actual number of verb types for each mood can be visualized in Table 6.5 and Table 6.6).

7 Conclusion and general discussion

model extends the previous ones, both in its architecture and in its explanatory power” (pp. 1-2). The studies presented aimed at contributing to this direction of research, by examining the implications of each theoretical proposal in detail and designing studies that directly test the divergent predictions. Only in this way can our understanding of the processes involved in morphological learning truly advance.

A general criticism towards the observations gathered through wug experiments with adult participants concerns their ecological validity. Proponents of the dual-mechanism model argue that children’s spontaneous linguistic behavior differs from that of adult speakers elicited in experimental settings. Children are observed to generalize irregular patterns by means of analogy less than adults (Pinker, 1995a). This concern can be addressed by repeating the experiment with children who are supposed to be more naïve about the objectives of the task. However, the linguistic behavior of the adult participants suggests that they are indeed able to store and track very fine-grained information about the distributional properties of the lexicon and use it to derive complex forms. This information goes beyond apparent similarities to high-frequency words. Indeed, rules extracted by the MGL model sometimes specified only a phoneme or phonetic features as the rule context. Such subtle patterns, if used by humans, are certainly below the level of conscious and strategic reasoning. Also, if a conscious similarity to a small number of highly salient verbs were at play, speakers would have responded using the morphological pattern of the same conjugation when deriving the infinitive and the

7 Conclusion and general discussion

subjunctive of the very same verb. Clearly, this was not the case²⁶. It is finally difficult to see how speakers would have consciously used a complex measure of type and token frequency combined to derive new forms. This performance points to properties of the brain that are beyond our active control.

The results regarding the production of subjunctive forms are especially interesting in light of the speech phenomena analyzed. Subjunctive morphology of the non-default conjugation classes proved to be the most productive in the experiment. This behavior is not exclusive to the experimental setting. We saw that in speech, oftentimes nonce subjunctive forms are produced. Speakers combine verbs with the incorrect suffix and derive non-existing forms. This indicates that, overall, the subjunctive is so infrequent, that it is hard to retrieve the correct pattern during articulation.

Overall, the facts concerning the use and productivity of the subjunctive gathered so far raise interesting questions about the processing of various subjunctive forms. Future lines of work must address to what extent affix productivity and other frequency measures impact processing duration using the subjunctive as a case study. This work will complement the study on the use, learning and generalization of the subjunctive mood presented in the present dissertation and,

²⁶ With the very same experiment stimuli, the participants used most frequently the pattern of the 1st conjugation in order to derive infinitive forms whereas they used the patterns of the other conjugations when deriving subjunctive forms (see again Figure 6.3).

7 Conclusion and general discussion

together with it, offer a more complete picture of how frequency shapes the most important aspects of language.

Appendix A: Corpus analysis

Model formula used to examine the influence of the governor on the mood distribution:

```
glmer(subjunctive~governor.verb+
      (1|embedded.verb)+(1|log.freq.s), data=d, na.action = na.omit,
      family=binomial, control=glmerControl(calc.derivs=F))
```

Appendix Table 1
Estimates of the generalized linear mixed model that predicts the occurrence of the subjunctive based on the governor

Governors	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.04330	0.45615	0.095	0.92437
<i>aspettare</i> “to wait”	-0.02415	1.69513	-0.014	0.98863
<i>bastare</i> “to suffice”	-1.15890	0.99145	-1.169	0.24245
<i>bisognare</i> “to necessitate”	3.45008	1.33580	2.583	0.00980 **
<i>credere</i> “believe”	1.13632	0.72525	1.567	0.11716
<i>in modo che</i> “so that”	0.76913	0.95562	0.805	0.42091
<i>nel caso</i> “in case”	0.63289	1.57062	0.403	0.68698
<i>non essere che</i> “to be not that”	-1.53346	0.62774	-2.443	0.01457 *

Appendix B: Experiment stimuli

Governors	Estimate	Std. Error	z value	Pr(> z)
<i>parere</i> “to seem”	2.40461	1.53776	1.564	0.11789
<i>pensare</i> “to think”	0.80495	0.66652	1.208	0.22717
<i>potere darsi</i> “it can be”	0.51016	0.97633	0.522	0.60130
<i>prima che</i> “before”	1.84138	1.07850	1.707	0.08776 .
<i>se</i> “if”	0.91192	0.56517	1.613	0.10663
<i>sembrare</i> “to seem”	1.16760	0.70685	1.652	0.09857 .
<i>sperare</i> “to hope”	1.88573	1.36958	1.377	0.16855
<i>volere</i> “to want”	3.42818	1.30940	2.618	0.884

Appendix B: Experiment stimuli

Appendix Table 2 shows the pseudoverbs used as stimuli in the experiment with their phonetic transcription. The verbs are ordered alphabetically. Overall 90 stimuli were used including four phonotactically illegal verbs: *zuisdo* [dzuizdo], *pebcro* [pebkro], *ghirilbo* [girilbo]. All the stimuli were presented to each of the 35 participants in randomized order. They were presented in the 1sg. present indicative form. The form has the same affix *-o* for all

Appendix B: Experiment stimuli

conjugations. Therefore, the participants could not tell beforehand to which conjugation the novel verb belongs. Transforming the pseudoverbs to generate an infinitive or a subjunctive form requires assigning the verb to one of the conjugation classes and choosing the appropriate morphological pattern.

Appendix Table 2
Stimuli used in the oral production task

Pseudoverbs	Phonetic transcription
<i>agguapro</i>	aggwapro
<i>areguo</i>	aregwo
<i>aremizzo</i>	aremiddzo
<i>artedo</i>	artædo
<i>atravo</i>	atravo
<i>arvinetto</i>	arvinnetto
<i>astumo</i>	astumo
<i>boglio</i>	boʎʎo
<i>bumo</i>	bumo
<i>burto</i>	burto
<i>carvesto</i>	karvesto
<i>catrido</i>	catrido
<i>cleniggo</i>	kleniggo
<i>coriundo</i>	korjundo
<i>crado</i>	krado
<i>crundo</i>	crundo

Appendix B: Experiment stimuli

Pseudoverbs	Phonetic transcription
<i>curpaio</i>	kurpajo
<i>drosumo</i>	drosumo
<i>erto</i>	erto
<i>ficido</i>	fifido
<i>folgo</i>	folgo
<i>fosvado</i>	fosvado
<i>frasto</i>	frasto
<i>frosco</i>	frosko
<i>fumpio</i>	fumpjo
<i>gaccio</i>	gattjo
<i>ghirilbo</i>	girilbo
<i>goglio</i>	gollo
<i>golvo</i>	golvo
<i>gresverto</i>	gresverto
<i>gringo</i>	gringo
<i>grobmo</i>	grobmo
<i>grolvo</i>	grolvo
<i>guisco</i>	gwisko
<i>iglisto</i>	il listino
<i>illipingo</i>	illipingo
<i>invoscedo</i>	invoscedo
<i>ircludo</i>	irkludo
<i>irnetto</i>	irnetto
<i>irniedo</i>	irnjedo

Appendix B: Experiment stimuli

Pseudoverbs	Phonetic transcription
<i>irquendo</i>	irkwendo
<i>isemo</i>	isɛmo
<i>liuco</i>	ljuko
<i>lotisco</i>	lotisko
<i>mico</i>	miko
<i>nieggo</i>	njɛggo
<i>nisdico</i>	nisdiko
<i>nistengo</i>	nistengo
<i>nuccedo</i>	nutʃɛdo
<i>ocoro</i>	okoro
<i>ostigo</i>	ostigo
<i>pebcro</i>	pebkro
<i>peguo</i>	pegwo
<i>pesfiggo</i>	pesfiggo
<i>pistengo</i>	pistengo
<i>pludo</i>	pludo
<i>rasulgo</i>	rasulgo
<i>rello</i>	rello
<i>rengo</i>	rengo
<i>resco</i>	resko
<i>reto</i>	reto
<i>rigovo</i>	rigovo
<i>rosongo</i>	rosongo
<i>rucuo</i>	rukwo

Appendix B: Experiment stimuli

Pseudoverbs	Phonetic transcription
<i>rungo</i>	rungo
<i>ruocio</i>	rwɔtʃo
<i>sado</i>	sado
<i>saio</i>	sajo
<i>sbergo</i>	sbergo
<i>sbossento</i>	sbossento
<i>scimango</i>	ʃimango
<i>screndo</i>	skrendo
<i>sgoto</i>	sgoto
<i>simo</i>	simo
<i>sisco</i>	sisko
<i>spelgo</i>	spelgo
<i>spurimo</i>	spurimo
<i>stietto</i>	stjetto
<i>striggio</i>	striddʒo
<i>tevo</i>	tevo
<i>tolvo</i>	tɔlvo
<i>topro</i>	topro
<i>tosisto</i>	tosisto
<i>trudo</i>	trudo
<i>tuscio</i>	tʊʃʃo
<i>ufficido</i>	uffiʃʃido
<i>virbatto</i>	virbatto
<i>vumpio</i>	vumpjo

Appendix C: Experimental design

Pseudoverbs	Phonetic transcription
<i>vuonisco</i>	vwonisko
<i>zuisdo</i>	dzuizdo

Appendix C: Experimental design

In the present appendix, I illustrate by means of an example how the stimuli are associated with the sentences depending on the experiment presentation and on the participant.

The ninety stimuli were divided in two groups: 45 were randomly assigned to the first group and the other 45 were assigned to the second group. The pseudoverb *simo* was randomly assigned to the first group, which in experiment presentation 1 appears together with infinitive sentences (see Appendix Table 3). The same verb appears together with subjunctive sentences in experiment presentation 2 (see).

Appendix Table 3 shows that the stimulus is randomly assigned to an input sentence and an infinitive test sentence for each participant. Thus, participant X, Carlo, sees the novel verb *simo* in input sentence A and must respond to infinitive sentence R while participant Y, Maria, sees the same stimulus in input sentence B and must respond to infinitive test sentence S.

Appendix Table 4 shows that stimulus *simo* is randomly assigned to an input sentence and a subjunctive test sentence for each participant: participant Z, Giorgio, receives the stimulus in input sentence C and must respond to subjunctive test sentence T. Participant Q, Eva,

Appendix C: Experimental design

receives the same stimulus in input sentence D and past respond to subjunctive test sentence U.

It is important to stress that the input sentence and the test sentence bear no relation to each other. First the test sentence is randomly assigned to the stimulus and then the input sentence is randomly assigned to the test sentence.

Appendix C: Experimental design

Appendix Table 3
Example of how two different participants may receive the stimulus *simo* in experiment presentation 1

Experiment presentation 1			
Participant	Stimulus	Random assignment of sentences to the stimulus	Example
Participant X “Carlo”	<i>simo</i>	Input sentence	<i>Spesso simo senza incertezze.</i>
		A	“Often I <i>simo</i> without hesitation.”
		Infinitive test sentence R	<i>Posso _____ quanto mi pare</i>
			“I can _____ as much as I want”
Participant Y “Maria”	<i>simo</i>	Input sentence	<i>Io simo prima di andarmene.</i>
		B	“I <i>simo</i> before I leave.”
		Infinitive test sentence S	<i>Non si può _____ senza avvisare.</i> “You cannot _____ without telling.”

Appendix C: Experimental design

Appendix Table 4
Example of how two different participants may receive the stimulus *simo* in experiment presentation 2

Experiment presentation 2			
Participant	Stimulus	Random assignment of sentences to the stimulus	Example
Participant Z “Giorgio”	<i>simo</i>	Input sentence C	Oggi simo se non ci siete. “Today I <i>simo</i> if you are not there.”
		Subjunctive test sentence T	Vuole che io _____ con loro. “He want that I _____ with them”
Participant Q “Eva”	<i>simo</i>	Input sentence D	Io simo con cautela “I <i>simo</i> with caution”
		Subjunctive test sentence U	Bisogna che io _____ presto “It’s necessary that I _____ soon”

Appendix C: Experimental design

The stimuli and their related sentences were presented in random order to each participant. The fillers were also presented in random order. The only constraint built in the experiment was that, after the appearance of the first filler, each subsequent filler appear approximately after 9 test sentences.

Appendix Table 5 shows the order in which one participant received the stimuli and the associated sentences in experiment presentation 1. Appendix Table 6 shows how another participant experienced experiment presentation 2. The verbs that occur together with infinitive test sentences in experiment presentation 1 are associated with subjunctive test sentences in experiment presentation 2 and vice-versa.

Let us take as an example pseudoverb *areguo*. It is randomly presented as the first stimulus in the example of experiment presentation 1 reported here (see order of presentation 1 in Appendix Table 5). It is randomly associated with an input sentence and a subjunctive test sentence.

The same pseudoverb randomly occurs as the ninth stimulus in the example of experiment presentation 2 (see order of presentation 9 in Appendix Table 6). It is randomly associated with an input sentence and an infinitive test sentence.

The two tables are not to be taken as a representation of how the input sentence and the test sentence appear on screen during the experiment. The input sentence together with the stimulus appears above, while the test sentence appears underneath it as in (C).

Appendix C: Experimental design

(C)

Ogni tanto **areguo** quell'impegno. (Input sentence)

Vuole che io _____ con fermezza. (Test sentence)

Appendix C: Experimental design

Appendix Table 5

Experiment presentation 1 as experienced by one participant

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
1	Ogni tanto areguo quell'impegno.	Vuole che io _____ con fermezza.	Subjunctive
2	Io ocoro perché non mi considerate.	Bisogna che io _____ mentre lo aspetto.	Subjunctive
3	Spesso resco senza incertezze.	Mi ha _____ volentieri.	Past participle (filler)
4	Io frasto sempre più velocemente.	Vuole che io _____ il giorno della vigilia.	Subjunctive
5	Io catrido in questo momento.	Vuole che io _____ con loro.	Subjunctive
6	Sono convinto che pistengo solo grazie a lui.	Non posso _____ al loro posto.	Infinitive
7	Oggi fumpio se non ci siete.	Posso _____ quanto mi pare.	Infinitive
8	Io crado prima di andarmene.	Ho imparato a _____ quando ero bambino.	Infinitive
9	Oggi ircludo da Parigi.	Bisogna che io _____ di sera tardi.	Subjunctive
10	Io nuccedo per non farla annoiare.	È stupido _____ senza farglielo sapere.	Infinitive
11	Non agguapro dopo cena.	Stamattina Maria ha _____ in fretta.	Past participle (filler)
12	Io grobmo una volta arrivato a casa.	Sono abituata a _____ in momenti difficili.	Infinitive
13	Non rasulgo con molta voglia.	Ora devi pensare a _____ .	Infinitive
14	Io burto dopo aver parlato con lei.	Vuole che io _____ .	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
15	Oggi tosisto molto prima.	Bisogna che io _____ presto.	Subjunctive
16	Io sbergo meno dei miei compagni.	Ha promesso di _____ senza fermarsi mai.	Infinitive
17	Non stietto mai con convinzione.	Bisogna che io _____ leggendo.	Subjunctive
18	Io bumo normalmente di sera.	Vuole che io _____ più in là.	Subjunctive
19	Io carvesto all'arrivo di mamma.	Marco ha _____ poco prima di te.	Past participle (filler)
20	Adesso nisdico senza pensarci troppo.	Vuole che io _____ sporadicamente.	Subjunctive
21	Non golvo di più degli altri.	Non ti permettere di _____ in cantina.	Infinitive
22	Io drosumo con la vostra presenza.	Vuole che io _____ mentre ti prepari.	Subjunctive
23	Non rungo se non me lo chiedete.	Ha potuto _____ anche da solo.	Infinitive
24	Io pesfiggo solo d'inverno.	Non si può _____ senza avvisare.	Infinitive
25	Non tuscio spesso in stazione.	Vuole che io _____ quando lei non c'è.	Subjunctive
26	Non ufficido per obbligo.	Ce la farà a _____ entro il termine previsto?	Infinitive
27	Oggi fumpio se non ci siete.	Ieri ho _____ con i miei amici.	Past participle (filler)
28	Io erto senza adeguarmi al resto.	Vuole che io _____ qualcosa di nuovo.	Subjunctive
29	Io non grolo per invidia.	Vuole che io _____ muovendomi con cautela.	Subjunctive
30	Io coriundo mentre papà guarda la TV.	Lascia lo spazio per _____.	Infinitive

Appendix C: Experimental design

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
31	Io mico per poco tempo.	Bisogna _____ a ogni costo.	Infinitive
32	Io tolvo volontariamente.	Lei ha l'esperienza per _____ a questi livelli.	Infinitive
33	Io irquendo anche per aiutarti.	Bisogna che io _____ una volta per tutte.	Subjunctive
34	Raramente zuisdo in questa sede.	Non ha ancora imparato a _____ .	Infinitive
35	Io irnetto con cautela.	L'altro giorno avete _____ senza avvisarmi.	Past participle (filler)
36	Io vumpio al suo posto.	Vuole che io _____ al suo posto.	Subjunctive
37	Io boglio moderatamente.	Bisogna che io _____ con ostinazione.	Subjunctive
38	Oggi rengo qui con voi.	Non è mio compito _____ anche per lei.	Infinitive
39	Non ne tevo più.	Ha un buon motivo per _____ in questo modo.	Infinitive
40	Io non fosvado per due ore.	Non sarò io a _____ per conto suo.	Infinitive
41	Sempre virbatto con ostinazione.	Vuole che io _____ per aiutarli.	Subjunctive
42	Io rello dopodomani.	L'ha vista _____ stamane.	Infinitive
43	Io gringo questi argomenti.	Voi avete _____ senza gli altri.	Past participle (filler)
44	Io ostigo di persona.	Vuole che io _____ con il bel tempo.	Subjunctive
45	Quando screndo sono felice.	Dobbiamo _____ con disinvoltura.	Infinitive
46	Io curpaio già da tempo.	Bisogna che io _____ ogni tanto con lui.	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
47	Io gringo questi argomenti.	Bisogna che io _____ spesso.	Subjunctive
48	È da ieri che sado invano.	Deve _____ in maniera appropriata.	Infinitive
49	Io ficido in solitudine.	Bisogna che io _____ con garbo.	Subjunctive
50	Non guisco da sole due settimane.	Vuole che io _____ per non dimenticare.	Subjunctive
51	Non nistengo durante la partita.	Non so se loro hanno _____ dopo di me.	Past participle (filler)
52	Non nistengo durante la partita.	Ti pare giusto _____ senza di me?	Infinitive
53	Spesso resco senza incertezze.	Mi sa che ha deciso di _____.	Infinitive
54	Io spurimo anche senza di te.	Bisogna che io _____ le giustificazioni degli altri.	Subjunctive
55	Io gaccio senza fare lamentele.	Vuole che io _____ con responsabilità.	Subjunctive
56	Io pebcro in poche occasioni.	Dovrebbe poter _____ anche lui.	Infinitive
57	Non saio nemmeno io.	Prevede di _____ per un bel po'.	Infinitive
58	Oggi artravo prima di tutti.	Vuole che io _____ per lui.	Subjunctive
59	Io pludo la comunità.	Giulia a Anna hanno _____ con molte remore.	Past participle (filler)
60	In tutti i casi non crundo da solo.	Vuole che io _____ da sola.	Subjunctive
61	Io spelgo verso le montagne.	Ha chiesto di _____ con noi.	Infinitive
62	Quasi sempre invoscedo con gli amici.	Bisogna che io _____ appena posso.	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
63	Io sisco dopo averlo visto.	Non ha previsto di _____ per tanti giorni.	Infinitive
64	Io artedo per passare il tempo.	Bisogna che io _____ da solo.	Subjunctive
65	Domani gresverto in due minuti.	Bisogna che io _____ alle quattro.	Subjunctive
66	Spesso io irniedo senza rendermene conto.	Bisogna che io _____ senza confondermi.	Subjunctive
67	Oggi artravo prima di tutti.	Poche ore fa ho _____ insieme a mio fratello.	Past participle (filler)
68	Io reto per un'oretta.	A che serve _____ in queste circostanze.	Infinitive
69	Spesso frosco a casa sua.	Bisogna che io _____ nuove cose.	Subjunctive
70	Sono due giorni che isemo .	Vuole che io _____ con poche persone.	Subjunctive
71	Io pludo la comunità.	Ha deciso di _____ prima.	Infinitive
72	Io irnetto con cautela.	Bisogna che io _____ dopodomani.	Subjunctive
73	Io rigovo per agevolarti.	Non ho mai capito come _____ questa roba.	Infinitive
74	Oggi ruocio perché c'è il sole.	Hanno potuto _____ in poco tempo.	Infinitive
75	Io rello dopodomani.	La settimana scorsa hanno _____ tardi.	Past participle (filler)
76	Oggi striggio contro voglia.	Bisogna che io _____ dopo aver parlato con lui.	Subjunctive
77	Io vuonisco meno degli altri.	Vai a _____ quello che vuoi.	Infinitive
78	Non agguapro dopo cena.	Vuole che io _____ dall'oggi al domani.	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
79	Io aremizzo da solo.	Bisogna che io _____ di nuovo.	Subjunctive
80	Io ghirilbo la rappresentanza.	Si ricorda spesso di _____ quando torna.	Infinitive
81	Io cleniggo anche domani.	Bisogna che io _____ le persone come te.	Subjunctive
82	Di pomeriggio liuco per poco.	Ha iniziato a _____ a sua insaputa.	Infinitive
83	Io simo dopo aver raccolto le critiche.	Maria vuole _____ a casa sua.	Infinitive
84	Non sbossento dopo di loro.	Bisogna che io _____ con pochi amici.	Subjunctive
85	Io topro all'accademia.	Vuole che io _____ aspettando il suo compleanno.	Subjunctive
86	Io folgo al posto loro.	Bisogna che io _____ prima di parlarvi.	Subjunctive
87	Io peguo con poca grinta.	È bene _____ così.	Infinitive
88	Io iglisto per me stessa.	Vuole che io _____ in estate.	Subjunctive
89	Qualche volta trudo in treno.	Vorrebbero _____ fuori casa.	Infinitive
90	Con loro nieggo con molta delicatezza.	Gli è dispiaciuto _____ senza di voi.	Infinitive
91	Io carvesto all'arrivo di mamma.	Vuole che io _____ sulle sue affermazioni.	Subjunctive
92	Non sgoto se non mi parlate di persona.	Cerca di _____ presto.	Infinitive
93	Io rucuo in poche ore.	È convinto di _____ nel modo giusto.	Infinitive
94	Non rosongo in molti casi.	Mi piace _____ quando sono in vacanza.	Infinitive

Appendix C: Experimental design

Experiment as experienced by participant A (experiment presentation 1)			
Num.	Input sentence	Test Sentence	Type of test sentence
95	A giugno arvinetto di nuovo.	Bisogna che io _____ dopo di loro.	Subjunctive
96	Quando illipingo Michela ride.	Vuole che io _____ senza perdere tempo.	Subjunctive
97	Io lotisco all'aperto.	Bisogna _____ con costanza.	Infinitive
98	Ogni tanto scimango anche io.	Non puoi _____ ogni giorno.	Infinitive
99	Io astumo al piano inferiore.	Non è facile _____ ogni cosa.	Infinitive
100	Io goglio quando mi pare.	Ha pensato bene di _____ prima di ricevere lamentele.	Infinitive

Appendix Table 6

Experiment presentation 2 as experienced by one participant

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
1	Io vuonisco al posto loro.	Vuole che io _____ da sola.	Subjunctive
2	Non ircludo spesso in stazione.	Non sarò io a _____ per conto suo.	Infinitive
3	Io boglio solo d'inverno.	Poche ore fa ho _____ insieme a mio fratello.	Past participle (filler)
4	Io stietto dopo aver parlato con lei.	Vai a _____ quello che vuoi.	Infinitive
5	Io tuscio dopodomani.	È bene _____ così.	Infinitive
6	Io carvesto moderatamente.	Ti pare giusto _____ senza di me?	Infinitive
7	Quando golvo Michela ride.	Vuole che io _____ con fermezza.	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
8	Io drosumo normalmente di sera.	Mi piace _____ quando sono in vacanza.	Infinitive
9	Oggi areguo qui con voi.	Deve _____ in maniera appropriata.	Infinitive
10	Io screndo per agevolarti.	Vuole che io _____ con il bel tempo.	Subjunctive
11	Non irnetto da sole due settimane.	Giulia a Anna hanno _____ con molte remore.	Past participle (filler)
12	Io nieggo per passare il tempo.	Bisogna che io _____ alle quattro.	Subjunctive
13	Io curpaio volontariamente.	Non ha ancora imparato a _____.	Infinitive
14	Oggi tolvo da Parigi.	Vuole che io _____ quando lei non c'è.	Subjunctive
15	Non ghirilbo con molta voglia.	Vuole che io _____ in estate.	Subjunctive
16	Io ufficido in questo momento.	Vuole che io _____ il giorno della vigilia.	Subjunctive
17	Sono convinto che tevo solo grazie a lui.	Vuole che io _____ aspettando il suo compleanno.	Subjunctive
18	Io frasto in solitudine.	Non si può _____ senza avvisare.	Infinitive
19	Io non sbergo per due ore.	Stamattina Maria ha _____ in fretta.	Past participle (filler)
20	Io trudo prima di andarmene.	Bisogna che io _____ spesso.	Subjunctive
21	Io irniedo mentre papà guarda la TV.	Ha un buon motivo per _____ in questo modo.	Infinitive
22	In tutti i casi non invoscedo da solo.	Lei ha l'esperienza per _____ a questi livelli.	Infinitive
23	Io pebcro al suo posto.	Vuole che io _____ sporadicamente.	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
24	Quasi sempre rello con gli amici.	Bisogna che io _____ nuove cose.	Subjunctive
25	Oggi virbatto perché c'è il sole.	Ce la farà a _____ entro il termine previsto?	Infinitive
26	Io sisco perché non mi considerate.	Vuole che io _____ per lui.	Subjunctive
27	Io nieggo per passare il tempo.	Non so se loro hanno _____ dopo di me.	Past participle (filler)
28	Io isemo questi argomenti.	Gli è dispiaciuto _____ senza di voi.	Infinitive
29	Adesso irquendo senza pensarci troppo.	Non ti permettere di _____ in cantina.	Infinitive
30	Raramente sgoto in questa sede.	Bisogna che io _____ da solo.	Subjunctive
31	È da ieri che rungo invano.	Bisogna che io _____ presto.	Subjunctive
32	Non scimango dopo di loro.	Bisogna che io _____ ogni tanto con lui.	Subjunctive
33	Oggi agguapro se non ci siete.	Non è facile _____ ogni cosa.	Infinitive
34	A giugno astumo di nuovo.	Bisogna che io _____ con pochi amici.	Subjunctive
35	Io stietto dopo aver parlato con lei.	Marco ha _____ poco prima di te.	Past participle (filler)
36	Oggi rengo prima di tutti.	Vuole che io _____ qualcosa di nuovo.	Subjunctive
37	Io nuccedo all'aperto.	Vuole che io _____ le donne presenti.	Subjunctive
38	Spesso io pludo senza rendermene conto.	Vuole che io _____ dall'oggi al domani.	Subjunctive
39	Non spelgo se non me lo chiedete.	Vuole che io _____ al suo posto.	Subjunctive

Appendix C: Experimental design

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
40	Oggi reto molto prima.	Vuole che io _____ più in là.	Subjunctive
41	Io folgo al piano inferiore.	L'ha vista _____ stamane.	Infinitive
42	Io burto anche per aiutarti.	Si ricorda spesso di _____ quando torna.	Infinitive
43	Io catrido con la vostra presenza.	L'altro giorno avete _____ senza avvisarmi.	Past participle (filler)
44	Io iglisto già da tempo.	Dovrebbe poter _____ anche lui.	Infinitive
45	Io spurimo senza adeguarmi al resto.	Posso _____ quanto mi pare.	Infinitive
46	Non aremizzo nemmeno io.	La richiesta di _____ è stata accettata.	Infinitive
47	Io liuco in poche occasioni.	Bisogna che io _____ le persone come te.	Subjunctive
48	Ogni tanto crado quell'impegno.	Bisogna che io _____ una volta per tutte.	Subjunctive
49	Io non sbergo per due ore.	Bisogna che io _____ con garbo.	Subjunctive
50	Io rucuo per non farla annoiare.	Bisogna che io _____ le giustificazioni degli altri.	Subjunctive
51	Io topro all'accademia.	La settimana scorsa hanno _____ tardi.	Past participle (filler)
52	Qualche volta nisdico in treno.	Bisogna _____ con costanza.	Infinitive
53	Non artedo per obbligo.	Non posso _____ al loro posto.	Infinitive
54	Io grolvo per me stessa.	A che serve _____ in queste circostanze.	Infinitive
55	Non bumo se non mi parlate di persona.	Non è mio compito _____ anche per lei.	Infinitive

Appendix C: Experimental design

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
56	Io saio senza fare lamentele.	Bisogna che io _____ dopodomani.	Subjunctive
57	Non rasulgo durante la partita.	Bisogna che io _____ di nuovo.	Subjunctive
58	Sono due giorni che frosco .	Ha promesso di _____ senza fermarsi mai.	Infinitive
59	Io ufficido in questo momento.	Voi avete _____ senza gli altri.	Past participle (filler)
60	Domani illipingo in due minuti.	Bisogna _____ a ogni costo.	Infinitive
61	Io rosongo con cautela.	Vuole che io _____ sulle sue affermazioni.	Subjunctive
62	Non simo mai con convinzione.	Bisogna che io _____ dopo aver parlato con lui.	Subjunctive
63	Io gresverto la rappresentanza.	Lascia lo spazio per _____ .	Infinitive
64	Io fosvado in poche ore.	Vuole che io _____ muovendomi con cautela.	Subjunctive
65	Io ostigo meno dei miei compagni.	Non ha previsto di _____ per tanti giorni.	Infinitive
66	Oggi vumpio contro voglia.	Ha chiesto di _____ con noi.	Infinitive
67	Qualche volta nisdico in treno.	Mi ha _____ volentieri.	Past participle (filler)
68	Io nistengo per poco tempo.	Vuole che io _____ per aiutarli.	Subjunctive
69	Io catrido con la vostra presenza.	Ha iniziato a _____ a sua insaputa.	Infinitive
70	Non zuisdo in molti casi.	Vuole che io _____ .	Subjunctive
71	Io topro all'accademia.	Sono abituata a _____ in momenti difficili.	Infinitive

Appendix C: Experimental design

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
72	Quando fumpio sono felice.	Vuole che io _____ con responsabilità.	Subjunctive
73	Io rigovo una volta arrivato a casa.	Bisogna che io _____ mentre lo aspetto.	Subjunctive
74	Io guisco anche domani.	Vorrebbero _____ fuori casa.	Infinitive
75	Non sado di più degli altri.	Ieri ho _____ con i miei amici.	Past participle (filler)
76	Io mico di persona.	Bisogna che io _____ di sera tardi.	Subjunctive
77	Io ruocio verso le montagne.	Bisogna che io _____ leggendo.	Subjunctive
78	Non gaccio dopo cena.	Non ho mai capito come _____ questa roba.	Infinitive
79	Io lotisco la comunità.	Bisogna che io _____ dopo di loro.	Subjunctive
80	Io striggio per un'oretta.	Cerca di _____ presto.	Infinitive
81	Io erto da solo.	Ho imparato a _____ quando ero bambino.	Infinitive
82	Io peguo dopo aver raccolto le critiche.	Vuole che io _____ per non dimenticare.	Subjunctive
83	Sempre ocoro con ostinazione.	È convinto di _____ nel modo giusto.	Infinitive
84	Non ne gringo più.	Dobbiamo _____ con disinvoltura.	Infinitive
85	Con loro tosisto con molta delicatezza.	Ha pensato bene di _____ prima di ricevere lamentele.	Infinitive
86	Spesso artravo a casa sua.	Ora devi pensare a _____.	Infinitive
87	Io ficido all'arrivo di mamma.	È stupido _____ senza farglielo sapere.	Infinitive
88	Io goglio sempre più velocemente.	Bisogna che io _____ appena posso.	Subjunctive

Appendix D: Collinearity between predictor variables

Experiment as experienced by participant B (Experiment presentation 2)			
Num.	Input sentence	Test sentence	Type of test sentence
89	Non irnetto da sole due settimane.	Mi sa che ha deciso di _____ .	Infinitive
90	Io cleniggo dopo averlo visto.	Ha deciso di _____ prima.	Infinitive
91	Io crundo meno degli altri.	Prevede di _____ per un bel po'.	Infinitive
92	Io arvinetto anche senza di te.	Maria vuole _____ a casa sua.	Infinitive
93	Ogni tanto sbossento anche io.	Hanno potuto _____ in poco tempo.	Infinitive
94	Io non coriundo per invidia.	Vuole che io _____ senza perdere tempo.	Subjunctive
95	Io pistengo con poca grinta.	Bisogna che io _____ prima di parlarvi.	Subjunctive
96	Spesso resco senza incertezze.	Vuole che io _____ con loro.	Subjunctive
97	Io boglio solo d'inverno.	Ha potuto _____ anche da solo.	Infinitive
98	Di pomeriggio pesfiggo per poco.	Bisogna che io _____ senza confondermi.	Subjunctive
99	Io grobmo quando mi pare.	Bisogna che io _____ con ostinazione.	Subjunctive
100	Non sado di più degli altri.	Vuole che io _____ mentre ti prepari.	Subjunctive

Appendix D: Collinearity between predictor variables

The MGL rule reliabilities based on type frequency and the MGL rule reliabilities based on token frequency appear to be collinear in the regression models. This means that, on average, the rule that applies to

Appendix D: Collinearity between predictor variables

many different verb types also appears with many verb tokens in speech. As a matter of fact, the higher the number of different verbs to which the morphological pattern applies (type frequency), the higher its absolute frequency of occurrence in speech (token frequency).

The correlation coefficients are reported in Appendix Table 7 and Appendix Table 8 for ease of consultation. The higher the correlation number is, i.e. the closer it approximates 1, the higher the correlation between two variables.

Appendix Table 7

Coefficients of correlation between MGL rule reliabilities based on type frequency and MGL rule reliabilities based on token frequency for the infinitives of each conjugation

1 st conjugation infinitive	0.716
2 nd conjugation infinitive	0.888
3 rd conjugation infinitive	0.918
4 th conjugation infinitive	0.943

Appendix Table 8

Coefficients of correlation between MGL rule reliabilities based on type frequency and MGL rule reliabilities based on token frequency for the subjunctives of each conjugation

1 st conjugation subjunctive	0.653
2 nd /3 rd /4 th conjugation subjunctive	0.814

Appendix E: MGL rule reliabilities

The experiment stimuli (phonetically transcribed) are entered into the MGL algorithm. The MGL applies the morphological rules learned from the lexicon to derive infinitives and subjunctives. The rules are made of a phonetic change in a given phonetic environment. The MGL gives a reliability value to each rule. The reliability reflects the probability that the rule will be applied to the stimulus.

Appendix Table 9 shows the infinitive forms (output form) generated by the MGL model, along with their reliability values. Appendix Table 10 shows the subjunctive forms (output form) generated by the MGL along with their reliability values. The output form with the highest reliability value wins out on the competing forms.

Appendix Table 9
MGL rules for the formation of infinitives with their reliabilities based on type frequency

ID num.	Stimuli (input form)		Infinitive (output form)	Phonetic change			Phonetic environment	MGL rule reliabilities based on type freq.
1	aggwapro	->	aggwaprère	o	->	ère	/X__	0.002
2	aggwapro	->	aggwaprire	o	->	ire	/Xapr__	0.57
3	aggwapro	->	aggwaprere	o	->	ere	/X__	0.082
4	aggwapro	->	aggwaprare	o	->	are	/Xr__	0.907
5	aregwo	->	aregwère	o	->	ère	/X__	0.002
6	aregwo	->	aregwire	o	->	ire	/Xegw__	0.648
7	aregwo	->	aregwere	o	->	ere	/Xw__	0.215
8	aregwo	->	aregware	o	->	are	/X__	0.754
9	aremiddzo	->	aremiddzère	o	->	ère	/X__	0.002

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
10	aremiddzo	->	aremiddzire	o	->	ire	/X__	0.011
11	aremiddzo	->	aremiddzere	o	->	ere	/X__	0.082
12	aremiddzo	->	aremiddzare	o	->	are	/Xz__	0.993
13	artedo	->	artedére	o	->	ére	/Xd__	0.028
14	artedo	->	artedire	o	->	ire	/X__	0.011
15	artedo	->	artedere	o	->	ere	/Xɛd__	0.941
16	artedo	->	artedare	o	->	are	/X__	0.754
17	arvinnetto	->	arvinnettére	o	->	ére	/X__	0.002
18	arvinnetto	->	arvinnettire	o	->	ire	/Xt__	0.034
19	arvinnetto	->	arvinnettere	o	->	ere	/Xtt__	0.302
20	arvinnetto	->	arvinnettare	o	->	are	/Xt__	0.827
21	astumo	->	astumére	o	->	ére	/X__	0.002
22	astumo	->	astumire	o	->	ire	/X__	0.011
23	astumo	->	astumere	o	->	ere	/Xum__	0.273
24	astumo	->	astumare	o	->	are	/Xm__	0.792
25	atravo	->	atravére	o	->	ére	/X__	0.002
26	atravo	->	atravire	o	->	ire	/X__	0.011
27	atravo	->	atravere	o	->	ere	/Xv__	0.3
28	atravo	->	atravare	o	->	are	/Xav__	0.872
29	βολλο	->	βολλέρε	o	->	ére	/X__	0.002
30	βολλο	->	βολλίρε	o	->	ire	/X__	0.011
31	βολλο	->	bolére	λλο	->	lére	/Xo__	0.215
32	βολλο	->	βολλέρε	o	->	ere	/X__	0.082
33	βολλο	->	βολλάρε	o	->	are	/X__	0.754
34	bumo	->	bumére	o	->	ére	/X__	0.002
35	bumo	->	bumire	o	->	ire	/X__	0.011
36	bumo	->	bumere	o	->	ere	/Xum__	0.273

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
37	bumo	->	bumare	o	->	are	/Xm__	0.792
38	burto	->	burtére	o	->	ére	/X__	0.002
39	burto	->	burtire	o	->	ire	/Xrt__	0.183
40	burto	->	burtere	o	->	ere	/Xt__	0.113
41	burto	->	burtare	o	->	are	/Xt__	0.827
42	catrido	->	catridére	o	->	ére	/Xd__	0.028
43	catrido	->	catridire	o	->	ire	/X__	0.011
44	catrido	->	catridere	o	->	ere	/Xtrid__	0.57
45	catrido	->	catridare	o	->	are	/X__	0.754
46	crundo	->	crundére	o	->	ére	/Xd__	0.028
47	crundo	->	crundire	o	->	ire	/X__	0.011
48	crundo	->	crundere	o	->	ere	/Xnd__	0.633
49	crundo	->	crundare	o	->	are	/X__	0.754
50	drosumo	->	drosumére	o	->	ére	/X__	0.002
51	drosumo	->	drosumire	o	->	ire	/X__	0.011
52	drosumo	->	drosumere	o	->	ere	/Xum__	0.273
53	drosumo	->	drosumare	o	->	are	/Xm__	0.792
54	dzuizdo	->	dzuizdére	o	->	ére	/Xd__	0.028
55	dzuizdo	->	dzuizdire	o	->	ire	/X__	0.011
56	dzuizdo	->	dzuizdere	o	->	ere	/Xd__	0.491
57	dzuizdo	->	dzuizdare	o	->	are	/X__	0.754
58	erto	->	ertére	o	->	ére	/X__	0.002
59	erto	->	ertire	o	->	ire	/Xrt__	0.183
60	erto	->	ertere	o	->	ere	/Xt__	0.113
61	erto	->	ertare	o	->	are	/Xt__	0.827
62	fiŋido	->	fiŋidére	o	->	ére	/Xd__	0.028
63	fiŋido	->	fiŋidire	o	->	ire	/X__	0.011

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
64	fiŋido	->	fiŋidere	o	->	ere	/Xŋid__	0.561
65	fiŋido	->	fiŋidare	o	->	are	/X__	0.754
66	folgo	->	foldgere	go	->	dʒer e	/Xl__	0.267
67	folgo	->	folére	go	->	ére	/X__	0.07
68	folgo	->	folire	go	->	ire	/X__	0.065
69	folgo	->	folgére	o	->	ére	/X__	0.002
70	folgo	->	foλλére	lgo	->	λλér e	/X__	0.3
71	folgo	->	folgire	o	->	ire	/X__	0.011
72	folgo	->	folgere	o	->	ere	/X__	0.082
73	folgo	->	folgare	o	->	are	/X__	0.754
74	fosvado	->	fosvadére	o	->	ére	/Xad__	0.216
75	fosvado	->	fosvadire	o	->	ire	/X__	0.011
76	fosvado	->	fosvadere	o	->	ere	/Xd__	0.491
77	fosvado	->	fosvadare	o	->	are	/X__	0.754
78	fosvado	->	fosandare	vad o	->	and are	/X__	0.254
79	frasto	->	frastére	o	->	ére	/X__	0.002
80	frasto	->	frastire	o	->	ire	/Xt__	0.034
81	frasto	->	frastere	o	->	ere	/Xt__	0.113
82	frasto	->	frastare	o	->	are	/Xast__	0.887
83	frosko	->	frosŋere	ko	->	ŋere	/X__	0.015
84	frosko	->	froskére	o	->	ére	/X__	0.002
85	frosko	->	froskire	o	->	ire	/X__	0.011
86	frosko	->	froŋŋere	sko	->	ŋŋer e	/X__	0.027
87	frosko	->	froskere	o	->	ere	/X__	0.082

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
88	frosko	->	froskare	o	->	are	/X__	0.754
89	fumpjo	->	fumpjére	o	->	ére	/X__	0.002
90	fumpjo	->	fumpjire	o	->	ire	/X__	0.011
91	fumpjo	->	fumpjere	o	->	ere	/Xmpj__	0.31
92	fumpjo	->	fumpjare	o	->	are	/Xj__	0.914
93	fumpjo	->	fumpire	jo	->	ire	/Xmp__	0.31
94	gattjo	->	gattjére	tjo	->	jére	/Xa__	0.087
95	gattjo	->	gattjére	o	->	ére	/X__	0.002
96	gattjo	->	gattjire	o	->	ire	/X__	0.011
97	gattjo	->	gattjere	o	->	ere	/X__	0.082
98	gattjo	->	gattjare	o	->	are	/X__	0.754
99	gattjo	->	gattjére	tjo	->	jére	/Xa__	0.087
100	girilbo	->	girilbére	o	->	ére	/X__	0.002
101	girilbo	->	girilbire	o	->	ire	/X__	0.011
102	girilbo	->	girilbere	o	->	ere	/X__	0.082
103	girilbo	->	girilbare	o	->	are	/Xb__	0.776
104	goλλo	->	goλλére	o	->	ére	/X__	0.002
105	goλλo	->	goλλire	o	->	ire	/X__	0.011
106	goλλo	->	golére	λλo	->	lére	/Xo__	0.215
107	goλλo	->	goλλere	o	->	ere	/X__	0.082
108	goλλo	->	goλλare	o	->	are	/X__	0.754
109	golvno	->	golvére	o	->	ére	/X__	0.002
110	golvno	->	golvire	o	->	ire	/X__	0.011
111	golvno	->	golvare	o	->	ere	/Xolv__	0.825
112	golvno	->	golvare	o	->	are	/X__	0.754
113	gresverto	->	gresvertére	o	->	ére	/X__	0.002
114	gresverto	->	gresvertire	o	->	ire	/Xvert__	0.852

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)	Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
115	gresverto	->	gresvertere	o -> ere	/Xt__	0.113
116	gresverto	->	gresvertare	o -> are	/Xt__	0.827
117	gringo	->	grindžere	go -> džer e	/Xin__	0.833
118	gringo	->	grinére	go -> ére	/Xn__	0.182
119	gringo	->	grinire	go -> ire	/X__	0.065
120	gringo	->	gringére	o -> ére	/X__	0.002
121	gringo	->	gringire	o -> ire	/X__	0.011
122	gringo	->	gringere	o -> ere	/X__	0.082
123	gringo	->	gringare	o -> are	/X__	0.754
124	grobmo	->	grobmére	o -> ére	/X__	0.002
125	grobmo	->	grobmire	o -> ire	/X__	0.011
126	grobmo	->	grobmere	o -> ere	/Xm__	0.142
127	grobmo	->	grobmare	o -> are	/Xm__	0.792
128	grólvo	->	grólvére	o -> ére	/X__	0.002
129	grólvo	->	grólvire	o -> ire	/X__	0.011
130	grólvo	->	grólvare	o -> ere	/Xólv__	0.825
131	grólvo	->	grólvare	o -> are	/X__	0.754
132	gwisko	->	gwistǰere	ko -> ǰere	/X__	0.015
133	gwisko	->	gwiskére	o -> ére	/X__	0.002
134	gwisko	->	gwire	sko -> re	/Xi__	0.991
135	gwisko	->	gwiskire	o -> ire	/X__	0.011
136	gwisko	->	gwiǰǰere	sko -> ǰǰer e	/X__	0.027
137	gwisko	->	gwiskere	o -> ere	/X__	0.082
138	gwisko	->	gwiskare	o -> are	/X__	0.754
139	illipingo	->	illipindžere	go -> džer e	/Xin__	0.833

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
140	illipingo	->	illipinére	go	->	ére	/Xn__	0.182
141	illipingo	->	illipinire	go	->	ire	/X__	0.065
142	illipingo	->	illipingére	o	->	ére	/X__	0.002
143	illipingo	->	illipingire	o	->	ire	/X__	0.011
144	illipingo	->	illipingere	o	->	ere	/X__	0.082
145	illipingo	->	illipingare	o	->	are	/X__	0.754
146	iǎǎisto	->	iǎǎistére	o	->	ére	/X__	0.002
147	iǎǎisto	->	iǎǎistire	o	->	ire	/Xt__	0.034
148	iǎǎisto	->	iǎǎistere	o	->	ere	/Xist__	0.383
149	iǎǎisto	->	iǎǎistare	o	->	are	/Xt__	0.827
150	invofedo	->	invofedére	o	->	ére	/Xd__	0.028
151	invofedo	->	invofedire	o	->	ire	/X__	0.011
152	invofedo	->	invofedere	o	->	ere	/Xd__	0.491
153	invofedo	->	invofedare	o	->	are	/X__	0.754
154	irkludo	->	irkludére	o	->	ére	/Xd__	0.028
155	irkludo	->	irkludire	o	->	ire	/X__	0.011
156	irkludo	->	irkludere	o	->	ere	/Xlud__	0.898
157	irkludo	->	irkludare	o	->	are	/X__	0.754
158	irkwendo	->	irkwendére	o	->	ére	/Xd__	0.028
159	irkwendo	->	irkwendire	o	->	ire	/X__	0.011
160	irkwendo	->	irkwendere	o	->	ere	/Xend__	0.974
161	irkwendo	->	irkwendare	o	->	are	/X__	0.754
162	imetto	->	imettére	o	->	ére	/X__	0.002
163	imetto	->	imettire	o	->	ire	/Xt__	0.034
164	imetto	->	imettere	o	->	ere	/Xett__	0.384
165	imetto	->	imettare	o	->	are	/Xt__	0.827
166	irnjedo	->	irnjedére	o	->	ére	/Xd__	0.028

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)	Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
167	irnjedo	->	irnjedire	o -> ire	/X__	0.011
168	irnjedo	->	irnjedere	o -> ere	/Xd__	0.491
169	irnjedo	->	irnjedare	o -> are	/X__	0.754
170	isemo	->	isemére	o -> ére	/X__	0.002
171	isemo	->	isemire	o -> ire	/X__	0.011
172	isemo	->	isemere	o -> ere	/Xem__	0.786
173	isemo	->	isemare	o -> are	/Xm__	0.792
174	karvesto	->	karvestére	o -> ére	/X__	0.002
175	karvesto	->	karvestire	o -> ire	/Xvest__	0.825
176	karvesto	->	karvestere	o -> ere	/Xt__	0.113
177	karvesto	->	karvestare	o -> are	/Xt__	0.827
178	kleniggo	->	klenigdžere	go -> džere	/X__	0.268
179	kleniggo	->	klenigére	go -> ére	/X__	0.07
180	kleniggo	->	klenigire	go -> ire	/X__	0.065
181	kleniggo	->	kleniggére	o -> ére	/X__	0.002
182	kleniggo	->	kleniggire	o -> ire	/X__	0.011
183	kleniggo	->	kleniggere	o -> ere	/X__	0.082
184	kleniggo	->	kleniggare	o -> are	/X__	0.754
185	kleniggo	->	kleniddžere	ggo -> ddžere	/Xi__	0.887
186	korjundo	->	korjundére	o -> ére	/Xd__	0.028
187	korjundo	->	korjundire	o -> ire	/X__	0.011
188	korjundo	->	korjundere	o -> ere	/Xnd__	0.633
189	korjundo	->	korjundare	o -> are	/X__	0.754
190	krado	->	kradére	o -> ére	/Xad__	0.216
191	krado	->	kradire	o -> ire	/X__	0.011
192	krado	->	kradere	o -> ere	/Xd__	0.491

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
193	krado	->	kradare	o	->	are /Xrad__	0.526
194	kurpajo	->	kurparire	jo	->	rire /Xpa__	0.607
195	kurpajo	->	kurpajére	o	->	ére /X__	0.002
196	kurpajo	->	kurpajire	o	->	ire /X__	0.011
197	kurpajo	->	kurpajere	o	->	ere /X__	0.082
198	kurpajo	->	kurpajare	o	->	are /Xj__	0.914
199	ljuko	->	ljufère	ko	->	ǵere /X__	0.015
200	ljuko	->	ljukére	o	->	ére /X__	0.002
201	ljuko	->	ljukire	o	->	ire /X__	0.011
202	ljuko	->	ljukere	o	->	ere /X__	0.082
203	ljuko	->	ljukare	o	->	are /Xuk__	0.718
204	lotisko	->	lotiǵere	ko	->	ǵere /X__	0.015
205	lotisko	->	lotiskére	o	->	ére /X__	0.002
206	lotisko	->	lotire	sko	->	re /Xi__	0.991
207	lotisko	->	lotiskire	o	->	ire /X__	0.011
208	lotisko	->	lotiǵǵere	sko	->	ǵǵere /X__	0.027
209	lotisko	->	lotiskere	o	->	ere /X__	0.082
210	lotisko	->	lotiskare	o	->	are /X__	0.754
211	miko	->	miǵere	ko	->	ǵere /X__	0.015
212	miko	->	mikére	o	->	ére /X__	0.002
213	miko	->	mikire	o	->	ire /X__	0.011
214	miko	->	mikere	o	->	ere /X__	0.082
215	miko	->	mikare	o	->	are /Xik__	0.912
216	nisdiko	->	nisiǵere	ko	->	ǵere /X__	0.015
217	nisdiko	->	nisdire	ko	->	re /Xdi__	0.313
218	nisdiko	->	nisdikére	o	->	ére /X__	0.002
219	nisdiko	->	nisdikire	o	->	ire /X__	0.011

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)	Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
220	nisdiko	->	nisdikere	o -> ere	/X__	0.082
221	nisdiko	->	nisdikare	o -> are	/Xik__	0.912
222	nistengo	->	nistendzere	go -> dʒere	/Xn__	0.432
223	nistengo	->	nistenére	go -> ére	/Xten__	0.916
224	nistengo	->	nistenire	go -> ire	/X__	0.065
225	nistengo	->	nistengére	o -> ére	/X__	0.002
226	nistengo	->	nistengire	o -> ire	/X__	0.011
227	nistengo	->	nistengere	o -> ere	/X__	0.082
228	nistengo	->	nistengare	o -> are	/X__	0.754
229	njeggo	->	njegdzere	go -> dʒere	/X__	0.268
230	njeggo	->	njegére	go -> ére	/X__	0.07
231	njeggo	->	njegire	go -> ire	/X__	0.065
232	njeggo	->	njeggére	o -> ére	/X__	0.002
233	njeggo	->	njeggire	o -> ire	/X__	0.011
234	njeggo	->	njeggere	o -> ere	/X__	0.082
235	njeggo	->	njeggare	o -> are	/X__	0.754
236	njeggo	->	njeddzere	ggo -> ddʒere	/Xɛ__	0.887
237	nutʃedo	->	nutʃedére	o -> ére	/Xd__	0.028
238	nutʃedo	->	nutʃedire	o -> ire	/X__	0.011
239	nutʃedo	->	nutʃedere	o -> ere	/Xed__	0.941
240	nutʃedo	->	nutʃedare	o -> are	/X__	0.754
241	okoro	->	okorére	o -> ére	/X__	0.002
242	okoro	->	okorire	o -> ire	/Xr__	0.024
243	okoro	->	okorere	o -> ere	/X__	0.082
244	okoro	->	okorare	o -> are	/Xor__	0.979

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
245	ostigo	->	ostidžere	go	->	džer e /Xi__	0.204
246	ostigo	->	ostiére	go	->	ére /X__	0.07
247	ostigo	->	ostiire	go	->	ire /X__	0.065
248	ostigo	->	ostigére	o	->	ére /X__	0.002
249	ostigo	->	ostigire	o	->	ire /X__	0.011
250	ostigo	->	ostigere	o	->	ere /X__	0.082
251	ostigo	->	ostigare	o	->	are /Xtig__	0.825
252	pebkro	->	pebkrére	o	->	ére /X__	0.002
253	pebkro	->	pebkrire	o	->	ire /Xr__	0.024
254	pebkro	->	pebkrere	o	->	ere /X__	0.082
255	pebkro	->	pebkrare	o	->	are /Xkr__	0.786
256	pegwo	->	pegwére	o	->	ére /X__	0.002
257	pegwo	->	pegwire	o	->	ire /Xegw__	0.648
258	pegwo	->	pegwere	o	->	ere /Xw__	0.215
259	pegwo	->	pegware	o	->	are /X__	0.754
260	pesfiggo	->	pesfigdžere	go	->	džer e /X__	0.268
261	pesfiggo	->	pesfigére	go	->	ére /X__	0.07
262	pesfiggo	->	pesfigire	go	->	ire /X__	0.065
263	pesfiggo	->	pesfiggére	o	->	ére /X__	0.002
264	pesfiggo	->	pesfiggire	o	->	ire /X__	0.011
265	pesfiggo	->	pesfiggere	o	->	ere /X__	0.082
266	pesfiggo	->	pesfiggare	o	->	are /X__	0.754
267	pesfiggo	->	pesfiddžere	ggo	->	ddžer re /Xi__	0.887
268	pistengo	->	pistendžere	go	->	džer e /Xn__	0.432

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
269	pistengo	->	pistenére	go	->	ére /Xten__	0.916
270	pistengo	->	pistenire	go	->	ire /X__	0.065
271	pistengo	->	pistengére	o	->	ére /X__	0.002
272	pistengo	->	pistengire	o	->	ire /X__	0.011
273	pistengo	->	pistengere	o	->	ere /X__	0.082
274	pistengo	->	pistengare	o	->	are /X__	0.754
275	pludo	->	pludére	o	->	ére /Xd__	0.028
276	pludo	->	pludire	o	->	ire /X__	0.011
277	pludo	->	pludere	o	->	ere /Xlud__	0.898
278	pludo	->	pludare	o	->	are /X__	0.754
279	rasulgo	->	rasulǵere	go	->	ǵer e /Xl__	0.267
280	rasulgo	->	rasulére	go	->	ére /X__	0.07
281	rasulgo	->	rasulire	go	->	ire /X__	0.065
282	rasulgo	->	rasulǵére	o	->	ére /X__	0.002
283	rasulgo	->	rasulǵere	lgo	->	ǵer e /X__	0.3
284	rasulgo	->	rasulgire	o	->	ire /X__	0.011
285	rasulgo	->	rasulgere	o	->	ere /X__	0.082
286	rasulgo	->	rasulgare	o	->	are /Xulg__	0.396
287	rengo	->	rendǵere	go	->	ǵer e /Xn__	0.432
288	rengo	->	renére	go	->	ére /Xn__	0.182
289	rengo	->	renire	go	->	ire /X__	0.065
290	rengo	->	rengére	o	->	ére /X__	0.002
291	rengo	->	rengire	o	->	ire /X__	0.011
292	rengo	->	rengere	o	->	ere /X__	0.082
293	rengo	->	rengare	o	->	are /X__	0.754

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
294	resko	->	resʃere	ko	->	ʃere	/X__	0.015
295	resko	->	reskére	o	->	ére	/X__	0.002
296	resko	->	reskire	o	->	ire	/X__	0.011
297	resko	->	reʃʃere	sko	->	ʃʃere	/X__	0.027
298	resko	->	ruʃʃire	esk o	->	uʃʃir e	/X__	0.18
299	resko	->	reskere	o	->	ere	/X__	0.082
300	resko	->	reskare	o	->	are	/X__	0.754
301	reto	->	retére	o	->	ére	/X__	0.002
302	reto	->	retire	o	->	ire	/Xt__	0.034
303	reto	->	retere	o	->	ere	/Xt__	0.113
304	reto	->	retare	o	->	are	/Xet__	0.916
305	rello	->	rellére	o	->	ére	/X__	0.002
306	rello	->	rellire	o	->	ire	/X__	0.011
307	rello	->	rellere	o	->	ere	/Xell__	0.57
308	rello	->	rellare	o	->	are	/Xl__	0.976
309	rigovo	->	rigovére	o	->	ére	/X__	0.002
310	rigovo	->	rigovire	o	->	ire	/X__	0.011
311	rigovo	->	rigovere	o	->	ere	/Xov__	0.852
312	rigovo	->	rigovare	o	->	are	/X__	0.754
313	rosongo	->	rosondžere	go	->	džere	/Xn__	0.432
314	rosongo	->	rosonére	go	->	ére	/Xn__	0.182
315	rosongo	->	rosonire	go	->	ire	/X__	0.065
316	rosongo	->	rosongére	o	->	ére	/X__	0.002
317	rosongo	->	rosongire	o	->	ire	/X__	0.011
318	rosongo	->	rosongere	o	->	ere	/X__	0.082

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
319	rosongo	->	rosongare	o	->	are /X__	0.754
320	rukwo	->	rukwére	o	->	ére /X__	0.002
321	rukwo	->	rukwire	o	->	ire /X__	0.011
322	rukwo	->	rukwere	o	->	ere /Xw__	0.215
323	rukwo	->	rukware	o	->	are /X__	0.754
324	rungo	->	rundzere	go	->	dzer e /Xun__	0.724
325	rungo	->	runére	go	->	ére /Xn__	0.182
326	rungo	->	runire	go	->	ire /X__	0.065
327	rungo	->	rungére	o	->	ére /X__	0.002
328	rungo	->	rungire	o	->	ire /X__	0.011
329	rungo	->	rungere	o	->	ere /X__	0.082
330	rungo	->	rungare	o	->	are /X__	0.754
331	rwɔŋfo	->	rwɔŋfère	o	->	ére /X__	0.002
332	rwɔŋfo	->	rwɔŋfire	o	->	ire /X__	0.011
333	rwɔŋfo	->	rwɔŋfere	o	->	ere /Xwɔŋf__	0.718
334	rwɔŋfo	->	rwɔŋfare	o	->	are /X__	0.754
335	sado	->	sadére	o	->	ére /Xad__	0.216
336	sado	->	sadire	o	->	ire /X__	0.011
337	sado	->	sadere	o	->	ere /Xd__	0.491
338	sado	->	sadare	o	->	are /X__	0.754
339	sajo	->	sajére	o	->	ére /X__	0.002
340	sajo	->	sajire	o	->	ire /X__	0.011
341	sajo	->	sajere	o	->	ere /X__	0.082
342	sajo	->	sajare	o	->	are /Xj__	0.914
343	sbergo	->	sberdzere	go	->	dzer e /Xer__	0.887
344	sbergo	->	sberére	go	->	ére /X__	0.07

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
345	sbergo	->	sberire	go	->	ire	/X__	0.065
346	sbergo	->	sbergére	o	->	ére	/X__	0.002
347	sbergo	->	sbergire	o	->	ire	/X__	0.011
348	sbergo	->	sbergere	o	->	ere	/X__	0.082
349	sbergo	->	sbergare	o	->	are	/X__	0.754
350	sbossento	->	sbossentére	o	->	ére	/X__	0.002
351	sbossento	->	sbossentire	o	->	ire	/Xsent__	0.786
352	sbossento	->	sbossentere	o	->	ere	/Xt__	0.113
353	sbossento	->	sbossentare	o	->	are	/Xnt__	0.923
354	sgoto	->	sgotére	o	->	ére	/X__	0.002
355	sgoto	->	sgotire	o	->	ire	/Xt__	0.034
356	sgoto	->	sgotere	o	->	ere	/Xt__	0.113
357	sgoto	->	sgotare	o	->	are	/Xot__	0.941
358	simo	->	simére	o	->	ére	/X__	0.002
359	simo	->	simire	o	->	ire	/X__	0.011
360	simo	->	simere	o	->	ere	/Xim__	0.28
361	simo	->	simare	o	->	are	/Xm__	0.792
362	sisko	->	sisʃere	ko	->	ʃere	/X__	0.015
363	sisko	->	siskére	o	->	ére	/X__	0.002
364	sisko	->	sire	sko	->	re	/Xi__	0.991
365	sisko	->	siskire	o	->	ire	/X__	0.011
366	sisko	->	siʃʃere	sko	->	ʃʃere	/X__	0.027
367	sisko	->	siskere	o	->	ere	/X__	0.082
368	sisko	->	siskare	o	->	are	/X__	0.754
369	skrendo	->	skrendére	o	->	ére	/Xd__	0.028
370	skrendo	->	skrendire	o	->	ire	/X__	0.011
371	skrendo	->	skrendere	o	->	ere	/Xnd__	0.633

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
372	skrendo	->	skrendare	o	->	are	/X__	0.754
373	spelgo	->	speldžere	go	->	džer e	/Xl__	0.267
374	spelgo	->	spelére	go	->	ére	/X__	0.07
375	spelgo	->	spelire	go	->	ire	/X__	0.065
376	spelgo	->	spelgére	o	->	ére	/X__	0.002
377	spelgo	->	spełłere	lgo	->	łłer e	/X__	0.3
378	spelgo	->	spelgire	o	->	ire	/X__	0.011
379	spelgo	->	spelgere	o	->	ere	/X__	0.082
380	spelgo	->	spelgare	o	->	are	/X__	0.754
381	spurimo	->	spurimére	o	->	ére	/X__	0.002
382	spurimo	->	spurimire	o	->	ire	/X__	0.011
383	spurimo	->	spurimere	o	->	ere	/Xrim__	0.765
384	spurimo	->	spurimare	o	->	are	/Xm__	0.792
385	stjetto	->	stjettére	o	->	ére	/X__	0.002
386	stjetto	->	stjettire	o	->	ire	/Xt__	0.034
387	stjetto	->	stjetttere	o	->	ere	/Xett__	0.384
388	stjetto	->	stjettare	o	->	are	/Xjett__	0.718
389	striddžo	->	striddžére	o	->	ére	/X__	0.002
390	striddžo	->	striddžire	o	->	ire	/X__	0.011
391	striddžo	->	striddžere	o	->	ere	/Xdž__	0.046
392	striddžo	->	striddžare	o	->	are	/Xddž__	0.947
393	ſimango	->	ſimandžere	go	->	džer e	/Xn__	0.432
394	ſimango	->	ſimanére	go	->	ére	/Xman__	0.57
395	ſimango	->	ſimanire	go	->	ire	/X__	0.065
396	ſimango	->	ſimangére	o	->	ére	/X__	0.002

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
397	ɟimango	->	ɟimangire	o	->	ire	/X__	0.011
398	ɟimango	->	ɟimangere	o	->	ere	/X__	0.082
399	ɟimango	->	ɟimangare	o	->	are	/X__	0.754
400	tevo	->	tevére	o	->	ére	/X__	0.002
401	tevo	->	tevíre	o	->	ire	/X__	0.011
402	tevo	->	tevere	o	->	ere	/Xv__	0.3
403	tevo	->	tevere	o	->	are	/X__	0.754
404	topro	->	toprére	o	->	ére	/X__	0.002
405	topro	->	topríre	o	->	ire	/Xpr__	0.61
406	topro	->	toprere	o	->	ere	/X__	0.082
407	topro	->	toprare	o	->	are	/Xr__	0.907
408	tosisto	->	tosistére	o	->	ére	/X__	0.002
409	tosisto	->	tosistire	o	->	ire	/Xt__	0.034
410	tosisto	->	tosistere	o	->	ere	/Xsist__	0.872
411	tosisto	->	tosistare	o	->	are	/Xt__	0.827
412	tólvo	->	tólvére	o	->	ére	/X__	0.002
413	tólvo	->	tólvire	o	->	ire	/X__	0.011
414	tólvo	->	tólvere	o	->	ere	/Xolv__	0.825
415	tólvo	->	tólvare	o	->	are	/X__	0.754
416	trudo	->	trudére	o	->	ére	/Xd__	0.028
417	trudo	->	trudire	o	->	ire	/X__	0.011
418	trudo	->	trudere	o	->	ere	/Xud__	0.813
419	trudo	->	trudare	o	->	are	/X__	0.754
420	tuʃʃo	->	tuʃʃére	o	->	ére	/X__	0.002
421	tuʃʃo	->	tuʃʃíre	o	->	ire	/X__	0.011
422	tuʃʃo	->	tuʃʃere	o	->	ere	/X__	0.082
423	tuʃʃo	->	tuʃʃare	o	->	are	/Xʃʃ__	0.833

Appendix E: MGL rule reliabilities

ID num.	Stimuli (input form)		Infinitive (output form)		Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
424	uffiʃido	->	uffiʃidére	o	->	ére	/Xd__	0.028
425	uffiʃido	->	uffiʃidire	o	->	ire	/X__	0.011
426	uffiʃido	->	uffiʃidere	o	->	ere	/Xʃid__	0.561
427	uffiʃido	->	uffiʃidare	o	->	are	/X__	0.754
428	virbatto	->	virbattére	o	->	ére	/X__	0.002
429	virbatto	->	virbattire	o	->	ire	/Xt__	0.034
430	virbatto	->	virbattere	o	->	ere	/Xbatt__	0.872
431	virbatto	->	virbattare	o	->	are	/Xt__	0.827
432	vumpjo	->	vumpjére	o	->	ére	/X__	0.002
433	vumpjo	->	vumpjire	o	->	ire	/X__	0.011
434	vumpjo	->	vumpjere	o	->	ere	/Xmpj__	0.31
435	vumpjo	->	vumpjare	o	->	are	/Xj__	0.914
436	vumpjo	->	vumpire	jo	->	ire	/Xmp__	0.31
437	vwonisko	->	vwonistjere	ko	->	tjere	/X__	0.015
438	vwonisko	->	vwoniskére	o	->	ére	/X__	0.002
439	vwonisko	->	vwonire	sko	->	re	/Xi__	0.991
440	vwonisko	->	vwoniskire	o	->	ire	/X__	0.011
441	vwonisko	->	vwonistjere	sko	->	tjere	/X__	0.027
442	vwonisko	->	vwoniskere	o	->	ere	/X__	0.082
443	vwonisko	->	vwoniskare	o	->	are	/X__	0.754

Appendix E: MGL rule reliabilities

Appendix Table 10

MGL rules for the formation of subjunctives with their reliabilities based on type frequency

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
1	aggwapro	->	aggwapri	o -> i	/Xr__	0.9
2	aggwapro	->	aggwapria	o -> ia	/X__	0.001
3	aggwapro	->	aggwapra	o -> a	/Xpr__	0.61
4	aggwapro	->	aggwaprki	o -> ki	/X__	0
5	aregwo	->	aregwi	o -> i	/X__	0.744
6	aregwo	->	aregwia	o -> ia	/X__	0.001
7	aregwo	->	aregwa	o -> a	/Xgw__	0.728
8	aregwo	->	aregwki	o -> ki	/X__	0
9	aremiddzo	->	aremiddzi	o -> i	/Xz__	0.984
10	aremiddzo	->	aremiddzia	o -> ia	/X__	0.001
11	aremiddzo	->	aremiddza	o -> a	/X__	0.236
12	aremiddzo	->	aremiddzki	o -> ki	/X__	0
13	artedo	->	artedi	o -> i	/X__	0.744
14	artedo	->	artedia	o -> ia	/X__	0.001
15	artedo	->	arteda	o -> a	/Xed__	0.941
16	artedo	->	artedki	o -> ki	/X__	0
17	arvinnetto	->	arvinnetti	o -> i	/Xt__	0.811
18	arvinnetto	->	arvinnettia	o -> ia	/X__	0.001
19	arvinnetto	->	arvinnetta	o -> a	/Xtt__	0.308
20	arvinnetto	->	arvinnettki	o -> ki	/X__	0
21	astumo	->	astumi	o -> i	/Xm__	0.785
22	astumo	->	astumia	o -> ia	/X__	0.001
23	astumo	->	astuma	o -> a	/Xum__	0.273
24	astumo	->	astumki	o -> ki	/X__	0
25	atravo	->	atravi	o -> i	/Xav__	0.852

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
26	atravo	->	atravia	o ->	ia	/X__	0.001
27	atravo	->	atrava	o ->	a	/Xv__	0.335
28	atravo	->	atravki	o ->	ki	/X__	0
29	boλλo	->	boλλi	o ->	i	/Xλλ__	0.917
30	boλλo	->	boλλia	o ->	ia	/X__	0.001
31	boλλo	->	boλλa	o ->	a	/X__	0.236
32	boλλo	->	boλλki	o ->	ki	/X__	0
33	bumo	->	bumi	o ->	i	/Xm__	0.785
34	bumo	->	bumia	o ->	ia	/X__	0.001
35	bumo	->	buma	o ->	a	/Xum__	0.273
36	bumo	->	bumki	o ->	ki	/X__	0
37	burto	->	burti	o ->	i	/Xt__	0.811
38	burto	->	burtia	o ->	ia	/X__	0.001
39	burto	->	burta	o ->	a	/Xrt__	0.254
40	burto	->	burtki	o ->	ki	/X__	0
41	catrido	->	catridi	o ->	i	/X__	0.744
42	catrido	->	catridia	o ->	ia	/X__	0.001
43	catrido	->	catrida	o ->	a	/Xtrid__	0.57
44	catrido	->	catridki	o ->	ki	/X__	0
45	crundo	->	crundi	o ->	i	/X__	0.744
46	crundo	->	crundia	o ->	ia	/X__	0.001
47	crundo	->	crunda	o ->	a	/Xnd__	0.629
48	crundo	->	crundki	o ->	ki	/X__	0
49	drosumo	->	drosumi	o ->	i	/Xm__	0.785
50	drosumo	->	drosunia	o ->	ia	/X__	0.001
51	drosumo	->	drosuna	o ->	a	/Xum__	0.273
52	drosumo	->	drosunki	o ->	ki	/X__	0
53	dzuizdo	->	dzuizdi	o ->	i	/X__	0.744

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
54	dzuizdo	->	dzuizdia	o	->	ia /X__	0.001
55	dzuizdo	->	dzuizda	o	->	a /Xd__	0.545
56	dzuizdo	->	dzuizdki	o	->	ki /X__	0
57	erto	->	erti	o	->	i /Xt__	0.811
58	erto	->	ertia	o	->	ia /X__	0.001
59	erto	->	erta	o	->	a /Xrt__	0.254
60	erto	->	ertki	o	->	ki /X__	0
61	fiʃido	->	fiʃidi	o	->	i /X__	0.744
62	fiʃido	->	fiʃidia	o	->	ia /X__	0.001
63	fiʃido	->	fiʃida	o	->	a /Xʃid__	0.561
64	fiʃido	->	fiʃidki	o	->	ki /X__	0
65	folgo	->	folgi	o	->	i /X__	0.744
66	folgo	->	folgia	o	->	ia /X__	0.001
67	folgo	->	folga	o	->	a /Xlg__	0.843
68	folgo	->	folgki	o	->	ki /X__	0
69	fosvado	->	fosvadi	o	->	i /X__	0.744
70	fosvado	->	fosvadia	o	->	ia /X__	0.001
71	fosvado	->	fosvada	o	->	a /Xd__	0.545
72	fosvado	->	fosvadki	o	->	ki /X__	0
73	frasto	->	frasti	o	->	i /Xast__	0.887
74	frasto	->	frastia	o	->	ia /Xst__	0.02
75	frasto	->	frasta	o	->	a /X__	0.236
76	frasto	->	frastki	o	->	ki /X__	0
77	frosko	->	froski	o	->	i /X__	0.744
78	frosko	->	froskia	o	->	ia /X__	0.001
79	frosko	->	froska	o	->	a /Xsk__	0.949
80	frosko	->	froskki	o	->	ki /X__	0
81	frosko	->	frosi	ko	->	i /X__	0.003

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
82	fumpjo	->	fumpji	o -> i	/Xj__		0.865
83	fumpjo	->	fumpjia	o -> ia	/X__		0.001
84	fumpjo	->	fumpja	o -> a	/Xmpj__		0.786
85	fumpjo	->	fumpjki	o -> ki	/X__		0
86	gatfjo	->	gatfji	o -> i	/X__		0.744
87	gatfjo	->	gatfjia	o -> ia	/X__		0.001
88	gatfjo	->	gatfja	o -> a	/Xatf__		0.363
89	gatfjo	->	gatfjki	o -> ki	/X__		0
90	girilbo	->	girilbi	o -> i	/Xb__		0.776
91	girilbo	->	girilbia	o -> ia	/X__		0.001
92	girilbo	->	girilba	o -> a	/X__		0.236
93	girilbo	->	girilbki	o -> ki	/X__		0
94	goλλo	->	goλλi	o -> i	/Xλλ__		0.917
95	goλλo	->	goλλia	o -> ia	/X__		0.001
96	goλλo	->	goλλa	o -> a	/X__		0.236
97	goλλo	->	goλλki	o -> ki	/X__		0
98	golv	->	golvi	o -> i	/X__		0.744
99	golv	->	golvia	o -> ia	/X__		0.001
100	golv	->	golva	o -> a	/Xolv__		0.825
101	golv	->	golvki	o -> ki	/X__		0
102	gresverto	->	gresverti	o -> i	/Xt__		0.811
103	gresverto	->	gresvertia	o -> ia	/X__		0.001
104	gresverto	->	gresverta	o -> a	/Xvert__		0.852
105	gresverto	->	gresvertki	o -> ki	/X__		0
106	gringo	->	gringi	o -> i	/X__		0.744
107	gringo	->	gringia	o -> ia	/X__		0.001
108	gringo	->	gringa	o -> a	/Xng__		0.89
109	gringo	->	gringki	o -> ki	/X__		0

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change	Phonetic environment	MGL rule reliabilities based on type freq.
110	grobmo	->	grobmi	o -> i	/Xm__	0.785
111	grobmo	->	grobmia	o -> ia	/X__	0.001
112	grobmo	->	grobma	o -> a	/X__	0.236
113	grobmo	->	grobmki	o -> ki	/X__	0
114	grølvo	->	grølvi	o -> i	/X__	0.744
115	grølvo	->	grølvia	o -> ia	/X__	0.001
116	grølvo	->	grølva	o -> a	/Xølv__	0.825
117	grølvo	->	grølvki	o -> ki	/X__	0
118	gwisko	->	gwiski	o -> i	/X__	0.744
119	gwisko	->	gwiskia	o -> ia	/X__	0.001
120	gwisko	->	gwiska	o -> a	/Xisk__	0.99
121	gwisko	->	gwiskki	o -> ki	/X__	0
122	gwisko	->	gwisi	ko -> i	/X__	0.003
123	illipingo	->	illipingi	o -> i	/X__	0.744
124	illipingo	->	illipingia	o -> ia	/X__	0.001
125	illipingo	->	illipinga	o -> a	/Xng__	0.89
126	illipingo	->	illipingki	o -> ki	/X__	0
127	iλλisto	->	iλλisti	o -> i	/Xt__	0.811
128	iλλisto	->	iλλistia	o -> ia	/Xst__	0.02
129	iλλisto	->	iλλista	o -> a	/Xist__	0.402
130	iλλisto	->	iλλistki	o -> ki	/X__	0
131	invoġedo	->	invoġedi	o -> i	/X__	0.744
132	invoġedo	->	invoġedia	o -> ia	/X__	0.001
133	invoġedo	->	invoġeda	o -> a	/Xd__	0.545
134	invoġedo	->	invoġedki	o -> ki	/X__	0
135	irkludo	->	irkludi	o -> i	/X__	0.744
136	irkludo	->	irkludia	o -> ia	/X__	0.001
137	irkludo	->	irkluda	o -> a	/Xlud__	0.898

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
138	irkludo	->	irkludki	o ->	ki	/X__	0
139	irkwendo	->	irkwendi	o ->	i	/X__	0.744
140	irkwendo	->	irkwendia	o ->	ia	/X__	0.001
141	irkwendo	->	irkwenda	o ->	a	/Xend__	0.973
142	irkwendo	->	irkwendki	o ->	ki	/X__	0
143	irnetto	->	irnetti	o ->	i	/Xt__	0.811
144	irnetto	->	irnettia	o ->	ia	/X__	0.001
145	irnetto	->	irnetta	o ->	a	/Xett__	0.375
146	irnetto	->	irnettki	o ->	ki	/X__	0
147	irnjedo	->	irnjedi	o ->	i	/X__	0.744
148	irnjedo	->	irnjedia	o ->	ia	/X__	0.001
149	irnjedo	->	irnjeda	o ->	a	/Xd__	0.545
150	irnjedo	->	irnjedki	o ->	ki	/X__	0
151	isemo	->	isemi	o ->	i	/Xm__	0.785
152	isemo	->	isemia	o ->	ia	/X__	0.001
153	isemo	->	isema	o ->	a	/Xem__	0.786
154	isemo	->	isemki	o ->	ki	/X__	0
155	karvesto	->	karvesti	o ->	i	/Xt__	0.811
156	karvesto	->	karvestia	o ->	ia	/Xst__	0.02
157	karvesto	->	karvesta	o ->	a	/Xvest__	0.825
158	karvesto	->	karvestki	o ->	ki	/X__	0
159	kleniggo	->	kleniggi	o ->	i	/X__	0.744
160	kleniggo	->	kleniggia	o ->	ia	/X__	0.001
161	kleniggo	->	klenigga	o ->	a	/Xgg__	0.955
162	kleniggo	->	kleniggki	o ->	ki	/X__	0
163	korjundo	->	korjundi	o ->	i	/X__	0.744
164	korjundo	->	korjundia	o ->	ia	/X__	0.001
165	korjundo	->	korjunda	o ->	a	/Xnd__	0.629

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
166	korjundo	->	korjundki	o	->	ki /X__	0
167	krado	->	kradi	o	->	i /X__	0.744
168	krado	->	kradia	o	->	ia /X__	0.001
169	krado	->	krada	o	->	a /Xd__	0.545
170	krado	->	kradki	o	->	ki /X__	0
171	kurpajo	->	kurpaji	o	->	i /Xj__	0.865
172	kurpajo	->	kurpajia	o	->	ia /X__	0.001
173	kurpajo	->	kurpaja	o	->	a /Xpaj__	0.872
174	kurpajo	->	kurpajki	o	->	ki /X__	0
175	ljuko	->	ljuki	o	->	i /X__	0.744
176	ljuko	->	ljukia	o	->	ia /X__	0.001
177	ljuko	->	ljuka	o	->	a /Xk__	0.533
178	ljuko	->	ljukki	o	->	ki /X__	0
179	ljuko	->	ljui	ko	->	i /X__	0.003
180	lotisko	->	lotiski	o	->	i /X__	0.744
181	lotisko	->	lotiskia	o	->	ia /X__	0.001
182	lotisko	->	lotiska	o	->	a /Xisk__	0.99
183	lotisko	->	lotiskki	o	->	ki /X__	0
184	lotisko	->	lotisi	ko	->	i /X__	0.003
185	miko	->	miki	o	->	i /Xik__	0.899
186	miko	->	mikia	o	->	ia /X__	0.001
187	miko	->	mika	o	->	a /Xk__	0.533
188	miko	->	mikki	o	->	ki /X__	0
189	miko	->	mii	ko	->	i /X__	0.003
190	nisdiko	->	nisdiki	o	->	i /Xik__	0.899
191	nisdiko	->	nisdikia	o	->	ia /X__	0.001
192	nisdiko	->	nisdika	o	->	a /Xk__	0.533
193	nisdiko	->	nisdikki	o	->	ki /X__	0

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
194	nisdiko	->	nisdii	ko	->	i /X__	0.003
195	nistengo	->	nistengi	o	->	i /X__	0.744
196	nistengo	->	nistengia	o	->	ia /X__	0.001
197	nistengo	->	nistenga	o	->	a /Xeng__	0.964
198	nistengo	->	nistengki	o	->	ki /X__	0
199	njeggo	->	njeggi	o	->	i /X__	0.744
200	njeggo	->	njeggia	o	->	ia /X__	0.001
201	njeggo	->	njegga	o	->	a /Xgg__	0.955
202	njeggo	->	njegcki	o	->	ki /X__	0
203	nutɸɛdo	->	nutɸɛdi	o	->	i /X__	0.744
204	nutɸɛdo	->	nutɸɛdia	o	->	ia /X__	0.001
205	nutɸɛdo	->	nutɸɛda	o	->	a /Xɛd__	0.941
206	nutɸɛdo	->	nutɸɛdki	o	->	ki /X__	0
207	okoro	->	okori	o	->	i /Xor__	0.978
208	okoro	->	okoria	o	->	ia /X__	0.001
209	okoro	->	okora	o	->	a /X__	0.236
210	okoro	->	okorki	o	->	ki /X__	0
211	ostigo	->	ostigi	o	->	i /Xtig__	0.825
212	ostigo	->	ostigia	o	->	ia /X__	0.001
213	ostigo	->	ostiga	o	->	a /Xg__	0.578
214	ostigo	->	ostigki	o	->	ki /X__	0
215	pebkro	->	pebkri	o	->	i /Xr__	0.9
216	pebkro	->	pebkria	o	->	ia /X__	0.001
217	pebkro	->	pebkra	o	->	a /X__	0.236
218	pebkro	->	pebkrki	o	->	ki /X__	0
219	pegwo	->	pegwi	o	->	i /X__	0.744
220	pegwo	->	pegwia	o	->	ia /X__	0.001
221	pegwo	->	pegwa	o	->	a /Xgw__	0.728

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
222	pegwo	->	pegwki	o	->	ki /X__	0
223	pesfiggo	->	pesfiggi	o	->	i /X__	0.744
224	pesfiggo	->	pesfiggia	o	->	ia /X__	0.001
225	pesfiggo	->	pesfigga	o	->	a /Xgg__	0.955
226	pesfiggo	->	pesfiggki	o	->	ki /X__	0
227	pistengo	->	pistengi	o	->	i /X__	0.744
228	pistengo	->	pistengia	o	->	ia /X__	0.001
229	pistengo	->	pistenga	o	->	a /Xeng__	0.964
230	pistengo	->	pistengki	o	->	ki /X__	0
231	pludo	->	pludi	o	->	i /X__	0.744
232	pludo	->	pludia	o	->	ia /X__	0.001
233	pludo	->	pluda	o	->	a /Xlud__	0.898
234	pludo	->	pludki	o	->	ki /X__	0
235	rasulgo	->	rasulgi	o	->	i /X__	0.744
236	rasulgo	->	rasulgia	o	->	ia /X__	0.001
237	rasulgo	->	rasulga	o	->	a /Xlg__	0.843
238	rasulgo	->	rasulgki	o	->	ki /X__	0
239	rengo	->	rengi	o	->	i /X__	0.744
240	rengo	->	rengia	o	->	ia /X__	0.001
241	rengo	->	renga	o	->	a /Xeng__	0.964
242	rengo	->	rengki	o	->	ki /X__	0
243	resko	->	reski	o	->	i /X__	0.744
244	resko	->	reskia	o	->	ia /X__	0.001
245	resko	->	reska	o	->	a /Xsk__	0.949
246	resko	->	reskki	o	->	ki /X__	0
247	resko	->	resi	ko	->	i /X__	0.003
248	reto	->	reti	o	->	i /Xet__	0.908
249	reto	->	retia	o	->	ia /X__	0.001

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
250	reto	->	reta	o ->	a	/X__	0.236
251	reto	->	retki	o ->	ki	/X__	0
252	reŋlo	->	reŋli	o ->	i	/Xl__	0.958
253	reŋlo	->	rellia	o ->	ia	/X__	0.001
254	reŋlo	->	rellā	o ->	a	/Xell__	0.57
255	reŋlo	->	rellki	o ->	ki	/X__	0
256	rigovo	->	rigovi	o ->	i	/X__	0.744
257	rigovo	->	rigovia	o ->	ia	/X__	0.001
258	rigovo	->	rigova	o ->	a	/Xov__	0.852
259	rigovo	->	rigovki	o ->	ki	/X__	0
260	rosongo	->	rosongi	o ->	i	/X__	0.744
261	rosongo	->	rosongia	o ->	ia	/X__	0.001
262	rosongo	->	rosonga	o ->	a	/Xng__	0.89
263	rosongo	->	rosongki	o ->	ki	/X__	0
264	rukwo	->	rukwi	o ->	i	/X__	0.744
265	rukwo	->	rukwia	o ->	ia	/X__	0.001
266	rukwo	->	rukwa	o ->	a	/Xw__	0.689
267	rukwo	->	rukwiki	o ->	ki	/X__	0
268	rungo	->	rungi	o ->	i	/X__	0.744
269	rungo	->	rungia	o ->	ia	/X__	0.001
270	rungo	->	runga	o ->	a	/Xng__	0.89
271	rungo	->	rungki	o ->	ki	/X__	0
272	rwɔŋfo	->	rwɔŋfi	o ->	i	/X__	0.744
273	rwɔŋfo	->	rwɔŋfia	o ->	ia	/X__	0.001
274	rwɔŋfo	->	rwɔŋfa	o ->	a	/Xwɔŋf__	0.718
275	rwɔŋfo	->	rwɔŋfki	o ->	ki	/X__	0
276	sado	->	sadi	o ->	i	/X__	0.744
277	sado	->	sadia	o ->	ia	/X__	0.001

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
278	sado	->	sada	o	->	a /Xd__	0.545
279	sado	->	sadki	o	->	ki /X__	0
280	sajo	->	saji	o	->	i /Xj__	0.865
281	sajo	->	sajia	o	->	ia /X__	0.001
282	sajo	->	saja	o	->	a /X__	0.236
283	sajo	->	sajki	o	->	ki /X__	0
284	sbergo	->	sbergi	o	->	i /X__	0.744
285	sbergo	->	sbergia	o	->	ia /X__	0.001
286	sbergo	->	sberga	o	->	a /Xerg__	0.887
287	sbergo	->	sbergki	o	->	ki /X__	0
288	sbossento	->	sbossenti	o	->	i /Xnt__	0.887
289	sbossento	->	sbossentia	o	->	ia /X__	0.001
290	sbossento	->	sbossenta	o	->	a /Xsent__	0.852
291	sbossento	->	sbossentki	o	->	ki /X__	0
292	sgoto	->	sgoti	o	->	i /Xot__	0.941
293	sgoto	->	sgotia	o	->	ia /X__	0.001
294	sgoto	->	sgota	o	->	a /X__	0.236
295	sgoto	->	sgotki	o	->	ki /X__	0
296	simo	->	simi	o	->	i /Xm__	0.785
297	simo	->	simia	o	->	ia /X__	0.001
298	simo	->	sima	o	->	a /Xim__	0.29
299	simo	->	simki	o	->	ki /X__	0
300	sisko	->	siski	o	->	i /X__	0.744
301	sisko	->	siskia	o	->	ia /X__	0.001
302	sisko	->	siska	o	->	a /Xisk__	0.99
303	sisko	->	siskki	o	->	ki /X__	0
304	sisko	->	sisi	ko	->	i /X__	0.003
305	skrendo	->	skrendi	o	->	i /X__	0.744

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
306	skrendo	->	skrendia	o ->	ia	/X__	0.001
307	skrendo	->	skrenda	o ->	a	/Xnd__	0.629
308	skrendo	->	skrendki	o ->	ki	/X__	0
309	spelgo	->	spelgi	o ->	i	/X__	0.744
310	spelgo	->	spelgia	o ->	ia	/X__	0.001
311	spelgo	->	spelga	o ->	a	/Xlg__	0.843
312	spelgo	->	spelgki	o ->	ki	/X__	0
313	spurimo	->	spurimi	o ->	i	/Xm__	0.785
314	spurimo	->	spurimia	o ->	ia	/X__	0.001
315	spurimo	->	spurima	o ->	a	/Xrim__	0.765
316	spurimo	->	spurimki	o ->	ki	/X__	0
317	stjetto	->	stjetti	o ->	i	/Xt__	0.811
318	stjetto	->	stjettia	o ->	ia	/X__	0.001
319	stjetto	->	stjetta	o ->	a	/Xett__	0.375
320	stjetto	->	stjettki	o ->	ki	/X__	0
321	striddʒo	->	striddʒi	o ->	i	/Xddʒ__	0.923
322	striddʒo	->	striddʒia	o ->	ia	/X__	0.001
323	striddʒo	->	striddʒa	o ->	a	/X__	0.236
324	striddʒo	->	striddʒki	o ->	ki	/X__	0
325	ʃimango	->	ʃimangi	o ->	i	/X__	0.744
326	ʃimango	->	ʃimangia	o ->	ia	/X__	0.001
327	ʃimango	->	ʃimanga	o ->	a	/Xng__	0.89
328	ʃimango	->	ʃimangki	o ->	ki	/X__	0
329	tevo	->	tevi	o ->	i	/X__	0.744
330	tevo	->	tevia	o ->	ia	/X__	0.001
331	tevo	->	teva	o ->	a	/Xv__	0.335
332	tevo	->	tevki	o ->	ki	/X__	0
333	topro	->	topri	o ->	i	/Xr__	0.9

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change		Phonetic environment	MGL rule reliabilities based on type freq.
334	topro	->	topria	o	-> ia	/X__	0.001
335	topro	->	topra	o	-> a	/Xpr__	0.61
336	topro	->	toprki	o	-> ki	/X__	0
337	tosisto	->	tosisti	o	-> i	/Xt__	0.811
338	tosisto	->	tosistia	o	-> ia	/Xst__	0.02
339	tosisto	->	tosista	o	-> a	/Xsist__	0.872
340	tosisto	->	tosistki	o	-> ki	/X__	0
341	tølvo	->	tølvi	o	-> i	/X__	0.744
342	tølvo	->	tølvia	o	-> ia	/X__	0.001
343	tølvo	->	tølva	o	-> a	/Xolv__	0.825
344	tølvo	->	tølvki	o	-> ki	/X__	0
345	trudo	->	trudi	o	-> i	/X__	0.744
346	trudo	->	trudia	o	-> ia	/X__	0.001
347	trudo	->	truda	o	-> a	/Xud__	0.823
348	trudo	->	trudki	o	-> ki	/X__	0
349	tuŋʃo	->	tuŋʃi	o	-> i	/Xʃʃ__	0.833
350	tuŋʃo	->	tuŋʃia	o	-> ia	/X__	0.001
351	tuŋʃo	->	tuŋʃa	o	-> a	/X__	0.236
352	tuŋʃo	->	tuŋʃki	o	-> ki	/X__	0
353	uffiŋʃido	->	uffiŋʃidi	o	-> i	/X__	0.744
354	uffiŋʃido	->	uffiŋʃidia	o	-> ia	/X__	0.001
355	uffiŋʃido	->	uffiŋʃida	o	-> a	/Xŋʃid__	0.561
356	uffiŋʃido	->	uffiŋʃidki	o	-> ki	/X__	0
357	virbatto	->	virbatti	o	-> i	/Xt__	0.811
358	virbatto	->	virbattia	o	-> ia	/X__	0.001
359	virbatto	->	virbatta	o	-> a	/Xbatt__	0.872
360	virbatto	->	virbattki	o	-> ki	/X__	0
361	vumpjo	->	vumpji	o	-> i	/Xj__	0.865

Appendix E: MGL rule reliabilities

ID num	Stimuli (input form)		Subjunctive (output form)	Phonetic change			Phonetic environment	MGL rule reliabilities based on type freq.
362	vumpjo	->	vumpjia	o	->	ia	/X__	0.001
363	vumpjo	->	vumpja	o	->	a	/Xmpj__	0.786
364	vumpjo	->	vumpjki	o	->	ki	/X__	0
365	vwonisko	->	vwoniski	o	->	i	/X__	0.744
366	vwonisko	->	vwoniskia	o	->	ia	/X__	0.001
367	vwonisko	->	vwoniska	o	->	a	/Xisk__	0.99
368	vwonisko	->	vwoniskki	o	->	ki	/X__	0
369	vwonisko	->	vwonisi	ko	->	i	/X__	0.003

References

- Abusch, D. (1997). Sequence of tense and temporal de re. *Linguistics and Philosophy*, 1–50.
- Adams, J. N. (2013). *Social variation and the Latin language*. Cambridge University Press.
- Adinolfi, G. (2017, February 4). *La lettera dei 600 docenti universitari al governo: “Molti studenti scrivono male, intervenite.”*
https://firenze.repubblica.it/cronaca/2017/02/04/news/firenze_la_lettera_dei_600_docenti_universitari_al_governo_molti_studenti_scrivono_male_intervenite_-157581214/
- Agha, A. (2006). *Language and social relations* (Vol. 24). Cambridge University Press.
- Albright, A. (2002a). Islands of reliability for regular morphology: Evidence from Italian. *Language*, 684–709.
- Albright, A. (2005). *The morphological basis of paradigm leveling*. na.
- Albright, A. (2008). Explaining universal tendencies and language particulars in analogical change. *Linguistic Universals and Language Change*, 144–181.

References

- Albright, A. (2009). Modeling analogy as probabilistic grammar. *Analogy in Grammar*, 3, 185–213.
- Albright, A. (2002b). The lexical bases of morphological well-formedness. *Morphology 2000: Selected Papers from the 9th Morphology Meeting*, 5–15.
- Albright, A., & Hayes, B. (2003). Rules vs. Analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2), 119–161.
- Aldai, G. (2000). Split ergativity in Basque: The pre-Basque antipassive-imperfective hypothesis. *Folia Linguistica Historica*, 34(Historica vol. 21, 1-2), 31–98.
- Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40(1), 41–61.
- Allaire, J. (2012). RStudio: Integrated development environment for R. *Boston, MA*, 537, 538.
- Allen, J. H., & Greenough, J. B. (2006). *New Latin grammar*. Courier Dover Publications.

References

- Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes*, 18(4), 405–430.
- Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, 3, 232.
- Anile, A. (1997). *Il cinema di Totò: (1930-1945): L'estro funambolo e l'ameno spettro*. Le Mani-Microart'S.
- Anile, A. (1998). *I film di Totò (1946-1967): La maschera tradita*. Mani.
- Anile, A. (2017). *Totalmente Totò. Vita e opere di un comico assoluto*. Cineteca di Bologna.
- Anshen, F., & Aronoff, M. (1988). Producing morphologically complex words. *Linguistics*, 26(4), 641–656.
- Antonelli, G. (2019). *Volgare eloquenza: Come le parole hanno paralizzato la politica*. Gius. Laterza & Figli Spa.
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67–82.

References

- Aronoff, M. (1976). Word formation in generative grammar. *Linguistic Inquiry Monographs Cambridge, Mass., 1*, 1–134.
- Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. MIT press.
- Baayen, H. (1992). Quantitative aspects of morphological productivity. In *Yearbook of morphology 1991* (pp. 109–149). Springer.
- Baayen, H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database (on CDROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Baayen, H., Schreuder, R., De Jong, N., & Krott, A. (2002). Dutch inflection: The rules that prove the exception. In *Storage and computation in the language faculty* (pp. 61–92). Springer.
- Baayen, R. H. (2003). Probabilistic approaches to morphology. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probability theory in linguistics* (pp. 229–287). Cambridge, Mass: The MIT Press.
- Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, 5(3), 436–461.

References

- Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37(1), 94–117.
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55(2), 290–313.
- Baayen, R. H., McQueen, J. M., Dijkstra, T., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. *Trends in Linguistics Studies and Monographs*, 151, 355–390.
- Baayen, R. H., Milin, P., \DJur\djević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118(3), 438.
- Baayen, R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon*, 2(3), 419–463.

References

- Bacelar do Nascimento, M., Casteleiro, J., Marques, M., Barreto, F., Amaro, R., & Veloso, R. (2000). *Léxico multifuncional computadorizado do português contemporâneo* [Computerized multifunctional lexicon of contemporary Portuguese]. *Lisboa: Centro de Linguística Da Universidade de Lisboa.*
- Baglioni, L. (2018, October 2). *Il congiuntivo*.
<https://www.youtube.com/watch?v=8bfYQZPLCEA>
- Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change*, 3(3), 241–264.
- Baldi, P., & Cuzzolin, P. (2010). *Constituent Syntax: Adverbial Phrases, Adverbs, Mood, Tense*. Walter de Gruyter.
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133(2), 283.
- Banniard, M. (2008). *Du latin aux langues romanes*. Armand Colin.
- Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009a). The WaCky wide web: A collection of very large linguistically

References

- processed web-crawled corpora. *Language Resources and Evaluation*, 43(3), 209–226.
- Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009b). The WaCky wide web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, 43(3), 209–226.
- Barr, R. C. (1995). *A lexical model of morphological change*. [PhD Thesis].
- Basile, N. (2001). *Sintassi storica del greco antico*. Levante.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*.
- Bates, D., Maechler, M., & Bolker, B. (2012). *Lme4: Linear mixed-effects models using S4 classes*. R package version 0.999999-0.
- Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R Package Version*, 2(1), 74.
- Beauvillain, C. (1996). The integration of morphological and whole-word form information during eye fixations on prefixed and

References

- suffixes. *Journal of Memory and Language*, 35(6), 801–820.
- Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92–111.
- Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., & Gildea, D. (1999). Forms Of English Function Words—Effects Of Disfluencies, Turn Position, Age And Sex, And Predictability. *IN PROCEEDINGS OF ICPHS-99*.
- Belsley, D., Kuh, E., & Welsch, R. (1980). Regression diagnostics Wiley. New York, NY.
- Benincà, P., Salvi, G., & Frison, L. (1988). L'ordine degli elementi della frase e le costruzioni marcate. *Grande Grammatica Italiana Di Consultazione*, 1, 115–225.
- Bennett, C. E. (1907). *The Latin language: A historical outline of its sounds, inflections, and syntax*. Allyn and Bacon.
- Bennett, C. E. (1946). *New Latin Grammar*. Allyn & Bacon.

References

- Benveniste, E. (1968). *Mutations of linguistic categories*. University of Texas Press.
- Bergareche, B. C. (1990). El futuro de subjuntivo en español. *Indicativo y Subjuntivo*, 410–427.
- Berko, J. (1958). The child's learning of English morphology. *Word*, 14(2–3), 150–177.
- Berretta, M. (1993). Morfologia. *Sobrero Alberto A. (a Cura Di), Introduzione All'italiano Contemporaneo, 1*, 193–245.
- Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, 42(3), 390–405.
- Biagi, M. L. A. (1987). *La grammatica dal testo: Grammatica italiana e testi per le scuole medie superiori*. Ugo Mursia Editore.
- Bianchi, V. (2001). *On Person Agreement*.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English* (Vol. 2). longman London.
- Bidois, G. L., & Bidois, R. L. (1971). *Syntaxe du français moderne*. II, Paris, Picard.

References

- Bispuri, E. (1997). *Totò: Principe clown: Tutti i film di Totò*. Guida Editori.
- Blake, R. (1981). Some empirically based observations on adult usage of the subjunctive mood in Mexico City. *Current Research in Romance Languages*, 13–22.
- Blevins, J. P. (2003). Stems and paradigms. *Language*, 79(4), 737–767.
- Blevins, J. P. (2006). English Inflection and Derivation. In B. Aarts & A. M. McMahon (Eds.), *The handbook of English linguistics* (pp. 507–536). Wiley Online Library.
- Blücher, K. (2003). Modalità, modo, "concordanza modale". Una prospettiva teorica. In M. Giacomo-Marcellesi & A. Rocchetti (Eds.), *Il verbo italiano* (pp. 1000–1010). Bulzoni.
- Blumenthal-Dramé, A. (2017). *Entrenchment from a psycholinguistic and neurolinguistic perspective*.
- Bolinger, D. (1968). Postposed main phrases: An English rule for the Romance subjunctive. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 14(1), 3–30.
- Bolinger, D. (1974). One subjunctive or two? *Hispania*, 57(3), 462–471.

References

- Bonfante, G. (1999). *“The” Origin of the Romance Languages: Stages in the Development of Latin*. C. Winter.
- Booij, G. E. (1977). *Dutch morphology: A study of word formation in generative grammar*. Peter de Ridder Press.
- Bopp, F., & Windischmann, K. J. (1816). *Franz Bopp über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache...* in der Andreäischen Buchhandlung.
- Bradley, D. (1980). Lexical representation of derivational relation. *Juncture*, 37–55.
- Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods*, 6(9), 2812–2831.
- Brysaert, (n.d.). *Dealing with zero word frequencies: A review of the existing rules of thumb and a suggestion for an evidence-based choice*.
- Brysaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure

References

- for American English. *Behavior Research Methods*, 41(4), 977–990.
- Bucholtz, M., & Hall, K. (2004). Language and identity. *A Companion to Linguistic Anthropology*, 1, 369–394.
- Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4–5), 585–614.
- Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language*, 18(4), 342–352.
- Burani, C., & Thornton, A. M. (2003). The interplay of root, suffix and whole-word frequency in processing derived words. *Trends in Linguistics Studies and Monographs*, 151, 157–208.
- Buratto, F. (2003). *Fantozzi una maschera italiana*. Lindau.
- Bush. (1999). *The predictive value of transitional probability for word boundary palatalization in English* [Master's Thesis]. University of New Mexico.
- Butragueño, P. M., & Lastra, Y. (2012). Corpus sociolingüístico de la ciudad de México. *Volumen II: Nivel Medio*. Mexico: El Colegio de México.

References

- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form* (Vol. 9). John Benjamins Publishing.
- Bybee, J. L. (1988). Morphology as lexical organization. *Theoretical Morphology*, 119141.
- Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5), 425–455.
- Bybee, J. L. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication*, 2(2), 249–278.
- Bybee, J. L. (2002a). Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. *Complex Sentences in Grammar and Discourse: Essays in Honor of Sandra A. Thompson*, 1–17.
- Bybee, J. L. (2002b). Sequentiality as the basis of constituent structure. *Typological Studies in Language*, 53, 109–134.
- Bybee, J. L. (2002c). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14(3), 261–290.
- Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge University Press.

References

- Bybee, J. L., & Brewer, M. A. (1980). Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua*, 52, 201–242.
- Bybee, J. L., & Hopper, P. J. (2001). *Frequency and the emergence of linguistic structure* (Vol. 45). John Benjamins Publishing.
- Bybee, J. L. & others. (2000). The phonology of the lexicon: Evidence from lexical diffusion. *Usage-Based Models of Language*, 65–85.
- Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world* (Vol. 196). University of Chicago Press Chicago.
- Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58(2), 265–289.
- Bybee, J. L., & Thompson, S. (1997). Three frequency effects in syntax. *Annual Meeting of the Berkeley Linguistics Society*, 23, 378–388.
- Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general

References

- principles of human cognition. *The Linguistic Review*, 22(2–4), 381–410.
- Calboli, G. (1969). *I modi del verbo greco e latino 1903-1966*. Vandenhoeck & Ruprecht.
- Campbell, L. (2001). What's wrong with grammaticalization? *Language Sciences*, 23, 113–162.
- Cappa, S., & Ullman, M. T. (1998). A neural dissociation in Italian verbal morphology. *5th Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA.[MTU]*.
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28(3), 297–332.
- Carlier, A., De Mulder, W., & Lamiroy, B. (2012). *Introduction: The pace of grammaticalization in a typological perspective*.
- Cesira, L. S. (2014). *La Sora Cesira*. <http://www.lasoracesira.it/>
- Cesira, L. S. (2018, December 18). *Il Congiuntivo*. <https://vimeo.com/307095923>
- Chafe, W. (1995). The realis-irrealis distinction in Caddo, the Northern Iroquoian languages, and English. *Modality in Grammar and Discourse*, 349–365.

References

- Chomsky, N. (1981). *1981: Lectures on government and binding*.
Dordrecht: Foris.
- Chomsky, N. (1995). *The minimalist program* Cambridge, MA: MIT Press.
- Chomsky, N. (1998). Minimalist inquiries: The framework. *MIT Occasional Papers in Linguistics*, 15.
- Chomsky, N. (2008). On phases. *Current Studies in Linguistics Series*, 45, 133.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*.
- Cicero, M. T. (1790). *Epistolarum selectarum libri IV* (I. V. der K. P. A. K. Buchhandlung, Ed.).
- Cicero, M. T. (1845). *Orationes: M. Tullii Ciceronis Oratio pro P. Sextio* (Vol. 1). Koehler.
- Cicero, M. T. (1856a). *Le opere di Marco Tullio Cicerone: Questioni Accademiche* (G. M. Scaramuzza, Ed.). G. Antonelli.
- Cicero, M. T. (1856b). *M. Tullii Ciceronis Orationes* (G. Long, Ed.; Vol. 3). Whittaker.
- Cicero, M. T. (1895). *M. Tulli Ciceronis Pro Publio Sestio: Oratio Ad Iudices* (H. A. Holden, Ed.). Macmillan and Company.

References

- Cicero, M. T. (1977). *Cicero: Epistulae Ad Familiares: Volume 2, 47-43 BC* (S. Bailey, Ed.). Cambridge University Press.
- Cicero, M. T., Clark, A. C., & Peterson, W. (1985). *Pro Sestio*. Buchner.
- Clahsen, H. (1995). German plurals in adult second language development. *The Current State of Interlanguage: Studies in Honor of William E. Rutherford*, 123–137.
- Clahsen, H. (1997). The representation of participles in the German mental lexicon: Evidence for the dual-mechanism model. In *Yearbook of morphology 1996* (pp. 73–95). Springer.
- Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22(6), 991–1013.
- Clahsen, H. (2006). *Dual-mechanism morphology*.
- Clahsen, H., Eisenbeiss, S., & Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguistics*, 23(3), 201–250.

References

- Clahsen, H., & Rothweiler, M. (1993). Inflectional rules in children's grammars: Evidence from German participles. In *Yearbook of morphology 1992* (pp. 1–34). Springer.
- Clahsen, H., Rothweiler, M., Woest, A., & Marcus, G. F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, 45(3), 225–255.
- Clancy, P. (1975). The rise of the synthetic future tense in Romance: A critique of some recent hypotheses (Benveniste, Valesio, Butler). *Romance Philology*, 28(4), 545–554.
- Cochran, W. G. (1952). The χ^2 test of goodness of fit. *The Annals of Mathematical Statistics*, 315–345.
- Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language*, 28(1), 1–13.
- Coleman, R. (1971). The origin and development of Latin habeo+infinitive. *The Classical Quarterly*, 21(1), 215–232.

References

Colombo, L., Stoianov, I., Pasini, M., & Zorzi, M. (2006). The role of phonology in the inflection of Italian verbs: A connectionist investigation. *The Mental Lexicon*, 1(1), 147–181.

Correggere mentalmente l'interlocutore con congiuntivi e condizionali. (2010, May 27).

[https://www.facebook.com/Correggere-mentalmente-linterlocutore-con-congiuntivi-e-condizionali-](https://www.facebook.com/Correggere-mentalmente-linterlocutore-con-congiuntivi-e-condizionali-114074421968667/?__tn__=%2Cd%2CP-R&eid=ARDUUkn_yWsAbOZQi3vB91IpjdnqhZwXYr6GL60r5XE7_BX8PIgJDYhc3P8KZK-XvqzYDlOvEdvcQoRW)

[114074421968667/?__tn__=%2Cd%2CP-](https://www.facebook.com/Correggere-mentalmente-linterlocutore-con-congiuntivi-e-condizionali-114074421968667/?__tn__=%2Cd%2CP-R&eid=ARDUUkn_yWsAbOZQi3vB91IpjdnqhZwXYr6GL60r5XE7_BX8PIgJDYhc3P8KZK-XvqzYDlOvEdvcQoRW)

[R&eid=ARDUUkn_yWsAbOZQi3vB91IpjdnqhZwXYr6GL6](https://www.facebook.com/Correggere-mentalmente-linterlocutore-con-congiuntivi-e-condizionali-114074421968667/?__tn__=%2Cd%2CP-R&eid=ARDUUkn_yWsAbOZQi3vB91IpjdnqhZwXYr6GL60r5XE7_BX8PIgJDYhc3P8KZK-XvqzYDlOvEdvcQoRW)

[0r5XE7_BX8PIgJDYhc3P8KZK-XvqzYDlOvEdvcQoRW](https://www.facebook.com/Correggere-mentalmente-linterlocutore-con-congiuntivi-e-condizionali-114074421968667/?__tn__=%2Cd%2CP-R&eid=ARDUUkn_yWsAbOZQi3vB91IpjdnqhZwXYr6GL60r5XE7_BX8PIgJDYhc3P8KZK-XvqzYDlOvEdvcQoRW)

Coseriu, E. (1981). Los conceptos de dialecto, nivel y estilo de lengua y el sentido propio de la dialectología. *LEA: Lingüística Española Actual*, 3(1), 1–32.

Costantini, F. (2014). Subject obviation in subjunctive clauses as a semantic failure. *Ms. University of Venice*.

Costantini, F. (2016a). Subject Obviation as a Semantic Failure: A Preliminary Account. *Annali Di Ca' Foscari. Serie Occidentale*, 50, 109–132.

References

- Costantini, F. (2016b). Subject obviation in subjunctive clauses as a semantic failure. *Annali Di Ca'Foscari. Serie Occidentale*, 50, 109–131.
- Cresti, E. (2000). *Corpus di italiano parlato: Introduzione* (Vol. 1). Accademia della Crusca.
- Cresti, E., & Moneglia, M. (2005). *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. John Benjamins, Amsterdam/Philadelphia.
- Cresti, E., Moneglia, M., do Nascimento, F. B., Moreno-Sandoval, A., Véronis, J., Martin, P., Choukri, K., Mapelli, V., Falavigna, D., Cid, A., & others. (2002). The C-ORAL-ROM Project. New methods for spoken language archives in a multilingual romance corpus. *LREC*.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Pearson Education.
- Croft, W., & Croft, W. A. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press on Demand.

References

- Crusca, A. della. (2014, March 27). *Accademia della Crusca*.
<https://www.facebook.com/AccademiaCrusca/photos/a.598007076909584/683671218343169/?type=1&theater>
- D'Achille, P. (2001). *Breve grammatica storica dell'italiano*. Carocci Rome.
- Dardano, M. (1978). *La formazione delle parole nell'italiano di oggi* (Vol. 148). Bulzoni.
- Dardano, M. (1983). Problemi della formazione delle parole. *Dardano, M., Dressler, W. e Held, G. (a Cura Di), Parallela, Tuebingen, Gunter Narr Verlag*, 9–25.
- Dardano, M., & Trifone, P. (1985). *La lingua italiana*. Zanichelli.
- Dardano, M., & Trifone, P. (1997). *La nuova grammatica della lingua italiana*. Zanichelli Editore.
- Daugherty, K. G., & Seidenberg, M. (1994). Beyond rules and exceptions: A connectionist approach to inflectional morphology. In R. L. Lima S. D. ., Corrigan & G. K. Iverson (Eds.), *The Reality of Linguistic Rules* (pp. 353–388). Amsterdam: John Benjamins.

References

- Davies, M. (2008). *The corpus of contemporary American English (COCA): 520 million words, 1990-present*.
- Davis, M. (2018). *Corpus del Español NOW*.
<https://www.corpusdelespanol.org/now/>
- De Cesare, A.-M. (2011). *L'ordine dei costituenti in italiano contemporaneo e in prospettiva contrastiva con il tedesco. Tra sintassi, pragmatica e tipologia linguistica* [PhD Thesis]. Habilitationsschrift, Istituto d'italianistica, Università di Basilea, MS.
- De Mauro, T., Mancini, F., Vedovelli, M., & Voghera, M. (1993). *Lessico di frequenza dell'italiano parlato*. Etas.
- De Mulder, W., & Lamiroy, B. (2012). Gradualness of grammaticalization in Romance. The position of French, Spanish and Italian. *Grammaticalization and Language Change: New Reflections*, 199–226.
- De Santis, C., & Fiorentino, G. (2018). La carica dei 600: La campagna mediatica sul declino della lingua italiana. *Circula: Revue d'idéologies Linguistiques*, 7, 1–29.

References

- De Santis, R. (2017, February 10). “*L’italiano non è in declino*”. *Il controappello dei linguisti alla lettera dei 600 accademici*.
https://www.repubblica.it/scuola/2017/02/10/news/_1_italiano_non_e_in_declino_il_controappello_dei_linguisti_alla_lettera_dei_600_accademici-157969348/
- De Sica, V. (1957). *L’oro di Napoli*. Paramount Pictures.
- De Smet, H. (2013). *Spreading patterns: Diffusional change in the English system of complementation*. Oxford University Press.
- Dee, J. H. (2002). The first downloadable word-frequency database for classical and medieval Latin. *The Classical Journal*, 98(1), 59–67.
- del Prado Martin, F. M., Kostić, A., & Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, 94(1), 1–18.
- Della Valle, V., & Patota, G. (2011). *Viva il congiuntivo! Come e quando usarlo senza sbagliare*. Sperling & Kupfer.
- Denis, D., & Sancier-Chateau, A. (1994). *Grammaire du français*. Librairie Générale Française.

References

- Denooz, J. (1978). L'ordinateur et le latin, Techniques et méthodes. *Revue de l'Organisation Internationale Pour l'Étude Des Langues Anciennes Par Ordinateur*, 4, 1–36.
- Denooz, J. (2004). Opera latina: Une base de données sur internet. *Euphrosyne*, 32, 79–88.
- Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience*, 10(1), 77–94.
- Diez, F. (1882). *Grammatik der romanischen Sprachen*. 5. Aufl. Eduard Weber's Verlag.
- Dion, N. (2003). L'effacement du que en français canadien: Une étude en temps réel. *MA Mémoire, University of Ottawa*.
- Distanziamento Sociale*. (2020, April 29). <https://www.facebook.com/photo.php?fbid=2588233928110541&set=gm.10157326639283108&type=3&theater>
- Dobrovie-Sorin, C. (2011). *The syntax of Romanian: Comparative studies in Romance* (Vol. 40). Walter de Gruyter.

References

- D'Onofrio: "Esperti, ditemi se è sbagliato il mio congiuntivo." (1994, September 27). La Repubblica. <https://ricerca.repubblica.it/repubblica/archivio/repubblica/1994/09/27/onofrio-esperti-ditemi-se.html>
- Dragosits, M. (2017). *Grammaticalization of the two Taboo Intensifiers damn and fucking in English* [PhD Thesis]. Universität Wien.
- Dressler, W. U., & Thornton, A. M. (1991). Doppie basi e binarismo nella morfologia italiana. *Rivista Di Linguistica*, 3(1), 3–22.
- Eckardt, R. (2012). Grammaticalization and semantic reanalysis. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning* (Vol. 3, pp. 2675–2702). Berlin : De Gruyter Mouton.
- Eckert, P. (2006). Communities of practice. *Encyclopedia of Language and Linguistics*, 2(2006), 683–685.
- Eckert, P. (2008). Variation and the indexical field 1. *Journal of Sociolinguistics*, 12(4), 453–476.
- Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, 41, 87–100.

References

- Eddington, D. (2002). Dissociation in Italian conjugations: A single-route account. *Brain and Language*, 81(1–3), 291–302.
- Ellis, N. C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The Modern Language Journal*, 92(2), 232–249.
- Ellis, N. C. (2012). What can we count in language, and what counts in language acquisition, cognition, and use. *Frequency Effects in Language Learning and Processing*, 1, 7–34.
- Farkas, D. (1992a). On obviation. *Lexical Matters*, 1, 85–109.
- Farkas, D. (1992b). On the semantics of subjunctive complements. *Romance Languages and Modern Linguistic Theory*, 69–104.
- Farkas, D. F. (1984). Subjunctive complements in Romanian. *Papers from the XIIIth Linguistic Symposium on Romance Languages, University Park, April 13, 1982*, 26, 355.
- Fazio, G. (2018, September 19). *Il congiuntivo è solo un modo verbale radical-chic che si può anche non usare?*
https://www.agi.it/cronaca/scuola_congiuntivo-4358564/news/2018-09-09/

References

- Ferrara, P. (2017, January 15). *L'agonia della lingua italiana in politica*. <http://www.pamela Ferrara.com/politici-congiuntivi-0115.html>
- Ferro, M., Marzi, C., & Pirrelli, V. (2018). Discriminative word learning is sensitive to inflectional entropy. *Lingue e Linguaggio*, 17(2), 307–327.
- Fidelholtz, J. L. (1975). Word frequency and vowel reduction in English. *Chicago Linguistic Society*, 11, 200–213.
- Firenze, G. di. (2007, February 4). *SAPER LEGGERE E SCRIVERE: UNA PROPOSTA CONTRO IL DECLINO DELL'ITALIANO A SCUOLA*. <https://www.edscuola.eu/wordpress/?p=86727>
- Fischer, J. L. (1958). Social influences on the choice of a linguistic variant. *Word*, 14(1), 47–56.
- Fisher, C., Gahl, S., Garnsey, S. M., & Matzen, L. (2006). That Sounds Unlikely: Syntactic Probabilities Affect Pronunciation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 28.
- Fo, D. (1995). *Totò: Manuale dell'attor comico*. Vallecchi.

References

- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1–2), 3–71.
- Francesco, I. (2018, February 16). *Di Maio, congiuntivo infinito: Ne sbaglia tre in una volta sola*.
<http://www.ilgiornale.it/news/politica/maio-congiuntivo-infinito-ne-sbaglia-tre-volta-sola-1495267.html>
- Gahl, S., & Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 748–775.
- García, M. E., & Terrell, T. (1977). Is the use of mood in Spanish subject to variable constraints. *Studies in Romance Linguistics*, 214–226.
- Garrett, A. (2008). Paradigmatic Uniformity and Markedness: : Historical convergence and universal grammar. In J. Good (Ed.), *Linguistic Universals and Language Change* (pp. 125–142). Oxford University Press.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, 12(5–6), 613–656.

References

- Geeslin, K. L., & Gudmestad, A. (2008). Comparing interview and written elicitation tasks in native and non-native data: Do speakers do what we think they do. *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, 64–77.
- Gentner, D., & Markman, A. B. (2002). *Analogy-based reasoning and metaphor*.
- Gentner, D., & Markman, A. B. (1998). Analogy-based reasoning. *The Handbook of Brain Theory and Neural Networks*, 91–93.
- Giannakidou, A. (1997). *The landscape of polarity items, unpublished Ph. D [PhD Thesis]*. dissertation, Groningen University.
- Gilquin, G., & Gries, S. T. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory*, 5(1), 1–26.
- Giorgi, A. (2009). Toward a syntax of the subjunctive mood. *Lingua*, 119(12), 1837–1858.
- Giorgi, A., & Pianesi, F. (1997). *Tense and aspect: From semantics to morphosyntax*. Oxford University Press on Demand.
- Giorgi, A., & Pianesi, F. (2000). Sequence of tense phenomena in Italian: A morphosyntactic analysis. *Probus*, 12(1), 1–32.

References

- Giorgi, A., & Pianesi, F. (2002). Sequence of Tense and the Speaker's Point of View: Evidence from the Imperfect. *Working Papers in Linguistics*, 12, 2002, Pp. 31-81.
- Giorgi, A., & Pianesi, F. (2004a). Complementizer deletion in Italian. *The Structure of CP and IP. The Cartography of Syntactic Structures*, 2, 190–210.
- Giorgi, A., & Pianesi, F. (2004b). The temporal perspective of the speaker and the subject: From semantics to morphosyntax. *The Syntax of Time*, 129–152.
- Giornale, I. (2018, November 25). *Premier-Fantozzi: Scivolone sul congiuntivo*. <http://www.ilgiornale.it/news/politica/premier-fantozzi-scivolone-sul-congiuntivo-1606780.html>
- Giraud, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15(4–5), 421–444.
- Giraud, H., Montermini, F., & Pirrelli, V. (2008). Processi cognitivi nell'analisi delle classi verbali dell'italiano: Un approccio

References

- sperimentale. *Atti Del XLII Congresso Della Società Di Linguistica Italiana*, IA-3.
- Givón, T. (1978). Negation in language: Pragmatics, function, ontology. *Syntax and Semantics*, 9(69), 105.
- Givón, T. (1987). Beyond foreground and background. *Coherence and Grounding in Discourse*, 11, 175–188.
- Givón, T. (1994). Irrealis and the subjunctive. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language,"* 18(2), 265–337.
- Givón, T., & Whitaker, H. A. (1979). *On understanding grammar* (Vol. 379). Academic Press.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press on Demand.
- Gonçalves, S. C. L. (2008). Projeto ALIP (Amostra Linguística do Interior Paulista): Questões teóricas e metodológicas sobre a constituição de um banco de dados de língua falada. *TAGNIN*,

References

- EO; VALE, OA *Avanços Da Linguística de Corpus No Brasil*.
São Paulo: Humanitas, 217–245.
- Graham, L. A. (2015). *The history of the future: Morphophonology, syntax, and grammaticalization* [PhD Thesis]. University of Georgia.
- Gramellini, M. (2007, June 13). *Basta che vi decidete*.
<https://www.lastampa.it/2007/06/13/cultura/basta-che-vi-decidete-Dcza8Ei9pcOyghjWSwPXJO/pagina.html?lgut=1>
- Grandgent, C. H. (1907). *An introduction to vulgar Latin*. DC Heath & Company.
- Grandgent, C. H. (1991). *Introducción al latín vulgar*. Editorial CSIC-CSIC Press.
- Greenberg, J. (1966). Language Universals. The Hague: Mouton. *Also in T. Sebeok (Ed.)*.
- Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society*, 35, 151–166.
- Grevisse, M. (1993). Le bon usage (by André Goosse). *Duculot, Paris*.

References

- Grossmann, M., & Rainer, F. (2013). *La formazione delle parole in italiano*. Walter de Gruyter.
- Gudmestad, A. (2010). Moving beyond a sentence-level analysis in the study of variable mood use in Spanish. *Southwest Journal of Linguistics*, 29(1).
- Gudmestad, A. (2012a). Acquiring a variable structure: An interlanguage analysis of second language mood use in Spanish. *Language Learning*, 62(2), 373–402.
- Gudmestad, A. (2012b). Toward an understanding of the relationship between mood use and form regularity: Evidence of variation across tasks, lexical items, and participant groups. *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, 214–227.
- Hahn, U., & Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, 41(4), 313–360.
- Haiman, J. (1985). Natural syntax. Iconicity and erosion. *Cambridge Studies in Linguistics London*, 44, 1–285.
- Haiman, J. (1991). From V/2 to subject clitics: Evidence from Northern Italian. *Approaches to Grammaticalization*, 2, 135–157.

References

- Haiman, J. (1994). Ritualization and the development of language. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, 3–3.
- Haiman, J. (2010). Competing motivations. In *The Oxford handbook of linguistic typology*.
- Haiman, J., & Thompson, S. A. (1988). *Clause combining in grammar and discourse* (Vol. 18). John Benjamins Publishing.
- Handford, S. A. (1947). *The Latin subjunctive: Its usage and development from Plautus to Tacitus*. Methuen & Company.
- Hare, M., & Elman, J. L. (1995). Learning and morphological change. *Cognition*, 56(1), 61–98.
- Hare, M., Elman, J. L., & Daugherty, K. G. (1995). Default generalisation in connectionist networks. *Language and Cognitive Processes*, 10(6), 601–630.
- Harris, M. (1974). The subjunctive mood as a changing category in Romance. *Historical Linguistics II: Theory and Description in Phonology*(*Proceedings of the First International Conference on Historical Linguistics, Edinburgh, 2-7, September, 1973*), 169–188.

References

- Harris, M. B. (1986). The historical development of conditional sentences in Romance. *Romance Philology*, 39(4), 405–436.
- Harwood, F. W., & Wright, A. M. (1956). Statistical study of English word formation. *Language*, 260–273.
- Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language,"* 22(2), 315–351.
- Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In *The new psychology of language* (pp. 217–248). Psychology Press.
- Heine, B. (1992). Grammaticalization chains. *Studies in Language. International Journal Sponsored by the Foundation "Foundations of Language,"* 16(2), 335–368.
- Heine, B. (1993). *Auxiliaries: Cognitive forces and grammaticalization*. Oxford University Press.
- Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction* (Vol. 9). Oxford University Press.
- Heine, B., Kuteva, T., & others. (2005). *Language contact and grammatical change*. Cambridge University Press.

References

- Heine, B., & Reh, M. (1984a). *Grammaticalization and reanalysis in African languages*. Buske Helmet Verlag GmbH.
- Heine, B., & Reh, M. (1984b). *Grammaticalization and reanalysis in African languages*. Buske Helmet Verlag GmbH.
- Herman, J. (2000). *Vulgar Latin*, translated by Roger Wright. University Park.
- Herzog, M. I., Labov, W., & Weinreich, U. (1968). Empirical foundations for a theory of language change. *WP Lehmann-Y. Malkiel (Hrsgg.), Directions for Historical Linguistics, Austin/London*.
- Higginbotham, J. (2002). Why is sequence of tense obligatory? *Logical Form and Language*, 207–227.
- Hock, H. (1991). *Principles of Historical Linguistics (2d ed.)*, Mouton de Gruyter. Berlin.
- Hodge, C. T. (1970). The linguistic cycle. *Language Sciences*, 13(7), 1–7.
- Hooper Bybee, J. (1979). Child morphology and morphophonemic change. *Linguistics*, 17(1–2), 21–50.

References

- Hooper, J. B. (1975). On Assertive Predicates in Syntax and Semantics, Vol. 4. *New York*.
- Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. *Current Progress in Historical Linguistics*, 96, 105.
- Hooper, J. B., & Terrell, T. (1974). A semantically based analysis of mood in Spanish. *Hispania*, 57(3), 484–494.
- Hooper, J. B., & Thompson, S. A. (1973). On the applicability of root transformations. *Linguistic Inquiry*, 4(4), 465–497.
- Hopper, P. J. & others. (1991). On some principles of grammaticization. *Approaches to Grammaticalization*, 1, 17–35.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge University Press.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417.
- Humboldt, W. von. (1822). Über das Entstehen der grammatischen Formen und ihren Einfluß auf die Ideenentwicklung.

References

- Abhandlungen Der Akademie Der Wissenschaften Zu Berlin*,
31–63.
- Iacobini, C. (2010). *Formazione delle Parole*. [http://www.treccani.it/enciclopedia/formazione-delle-parole_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/formazione-delle-parole_(Enciclopedia-dell'Italiano)/)
- Il congiuntivo non è una malattia degli occhi*. (2010, March 29).
<https://www.facebook.com/Il-congiuntivo-non-%C3%A8-una-malattia-degli-occhi-104014509638685/>
- Il congiuntivo non è una malattia degli occhi*. (2014, April 2).
<https://www.facebook.com/104014509638685/photos/rpp.104014509638685/747569961949800/?type=3&theater>
- Il congiuntivo non è una malattia degli occhi*. (2015, April 9).
<https://www.facebook.com/104014509638685/photos/rpp.104014509638685/960717080635086/?type=3&theater>
- Il Senato ha corretto un congiuntivo, in coro*. (2015, October 14). Il Post. <https://www.ilpost.it/2015/10/14/il-senato-ha-corretto-un-congiuntivo-in-coro/>

References

- Jaeggli, O., & Safir, K. J. (1989). The null subject parameter and parametric theory. In *The null subject parameter* (pp. 1–44). Springer.
- Jakobson, R. (1939). Signe zéro—Mélanges de linguistique offerts à Charles Bally. *R. Jakobson, Selected Writings II: Word and Language*, 211–219.
- Janson, T. (2004). *A natural history of Latin*. Oxford University Press.
- Jespersen, O. (1924). *The philosophy of grammar*. London, G. Allen & Unwin.
- Jezek, E. (2005). Lessico. Classi di parole, strutture, combinazioni. *II Mulino, Bologna*.
- Johnstone, B. (2016). Enregisterment: How linguistic items become linked with ways of speaking. *Language and Linguistics Compass*, 10(11), 632–643.
- Jones, S., & Sinclair, J. M. (2012). *English lexical collocations—A study in computational linguistics* (Vol. 1974). Classiques Garnier.
- Jurafsky, Dan. (2003). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. *Probabilistic Linguistics*, 21.

References

- Jurafsky, Daniel, Bell, A., Fosler-Lussier, E., Girand, C., & Raymond, W. (1998). Reduction of English function words in Switchboard. *Fifth International Conference on Spoken Language Processing*.
- Jurafsky, Daniel, Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language*, 45, 229–254.
- Jurafsky, Daniel, Bell, A., Gregory, M., & Raymond, W. D. (2002). The effect of language model probability on pronunciation reduction. *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, 2, 801–804.
- Kapatsinski, V. (2012). What statistics do learners track? Rules, constraints and schemas in (artificial) grammar learning. *Frequency Effects in Language Learning and Processing*, 1, 53–82.

References

- Kapatsinski, V., & Radicke, J. (2009). Frequency and the emergence of prefabs: Evidence from monitoring. *Formulaic Language*, 2, 499–520.
- Kay, P., & Fillmore, C. (1999). Construction grammar and linguistic generalizations: The *whats X doing Y?* construction. *Language*, 75, 1–33.
- Kempchinsky, P. (1986). *Romance Subjunctive Clauses and Logical Form*, unpublished Ph. D [PhD Thesis]. dissertation, UCLA.
- Kempchinsky, P. (1987). The subjunctive disjoint reference effect. *Studies in Romance Languages*, 25.
- Kempchinsky, P. (2009). What can the subjunctive disjoint reference effect tell us about the subjunctive? *Lingua*, 119(12), 1788–1810.
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, 48(1), 171.
- King, R. D. (1969). *Historical linguistics and generative grammar*. Prentice Hall.

References

- Kiparsky, P. (1968). Linguistic universals and linguistic change. In E. Bach & R. T. Harms (Eds.), *Universals in Linguistic Theory* (pp. 171–202). Holt, Rinehart, and Winston.
- Kiparsky, P. (1978). Analogical change as a problem for linguistic theory. *Studies in the Linguistic Sciences Urbana, Ill*, 8(2), 77–96.
- Klein Andreu, F. (1980). *A quantitative study of syntactic and pragmatic indications of change in the Spanish of bilinguals in the US*. In W. Labov (ed.), *Locating language in time and space*. 69–82. New York: Academic Press.
- Klein, F. (1975). Pragmatic constraints on distribution: The Spanish subjunctive. *11th CLS*, 353–365.
- Klein-Andreu, F. (1990). Losing ground a discourse-pragmatic solution to the history of -ra in Spanish. In S. Fleischman & L. R. Waugh (Eds.), *Discourse-pragmatics and the verb: Evidence from Romance* (pp. 164–178). Routledge.
- Kratzer, A. (1977). What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy*, 1(3), 337–355.

References

- Kratzer, A. (1979). Conditional necessity and possibility. In *Semantics from different points of view* (pp. 117–147). Springer.
- Kratzer, A., Eikmeyer, H.-J., & Rieser, H. (1981). The notional category of modality. *Formal Semantics: The Essential Readings*, 289–323.
- Krishnamurthy, R. (2006). Collocations. In K. Brown (Ed.), *Encyclopaedia of language and linguistics* (Vol. 2, pp. 596–600). Elsevier.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Dartmouth Publishing Group.
- Kuperman, V., Bertram, R., & Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, 62(2), 83–97.
- Kurylowicz, J. (1965). The evolution of grammatical categories. *Diogenes*, 13(51), 55–71.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19(3), 273–309.
- Labov, W. (1966). *The social stratification of English in New York city*.

References

- Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, 715–762.
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.
- Labov, W. (2001a). Principles of linguistic change Volume 2: Social factors. *LANGUAGE IN SOCIETY-OXFORD-*, 29.
- Labov, W. (2001b). Principles of linguistic change Volume 2: Social factors. *LANGUAGE IN SOCIETY-OXFORD-*, 29.
- Labov, W. (2011). *Principles of linguistic change, volume 3: Cognitive and cultural factors* (Vol. 36). John Wiley & Sons.
- Lafi, S., & Kaneene, J. (1992). An explanation of the use of principal-components analysis to detect and correct for multicollinearity. *Preventive Veterinary Medicine*, 13(4), 261–275.
- Laka, I. (1990). *Negation in syntax—on the nature of functional categories and projections* [PhD Thesis]. Massachusetts Inst. of Technology Cambridge.
- Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press.

References

- Lamiroy, B., & De Mulder, W. (2011). Degrees of grammaticalization across languages. *The Oxford Handbook of Grammaticalization*, 302–318.
- Landau, I. (2004). The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory*, 22(4), 811–877.
- Langacker, R. W. (1977). Syntactic reanalysis. *Mechanisms of Syntactic Change*, 59, 139.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford university press.
- Langacker, R. W. (2017). *Entrenchment in cognitive grammar*.
- Lauria, E. (2011, December 4). *Dieci candidati in campo, il big bang delle elezioni*.
https://palermo.repubblica.it/cronaca/2011/12/04/news/dieci_candidati_in_campo_il_big_bang_delle_elezioni-26056018/
- LeBlanc, C. L. (2009). Conditional morphology in si-clauses: A Canadian-French reanalysis. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 54(2), 317–337.

References

- Ledgeway, A. (2012). *From Latin to Romance: Morphosyntactic typology and change* (Vol. 1). Oxford University Press.
- Ledgeway, A., & Maiden, M. (2016). *The Oxford guide to the Romance languages* (Vol. 1). Oxford University Press.
- Leeman-Bouix, D. (1994). *Les fautes de français existent-elles?* FeniXX.
- Leggo. (2018, February 9). *Chi è Lorenzo Baglioni, “Il congiuntivo” è il video più visto tra i brani di Sanremo.* https://www.leggo.it/spettacoli/sanremo/lorenzo_baglioni_chi_e_videoclip_piu_visto_sanremo_9_febbraio_2018-3538103.html
- Lehmann, C. (1982). *Thoughts on grammaticalization: A programmatic sketch.* Universalienprojekt, Institut für Sprachwissenschaft Universität.
- Lehmann, C. (2015). *Thoughts on grammaticalization* (3rd ed.). Language Science Press.
- Leone, R. (2018, September 21). *Venghino signori pensionati, venghino al Sud.* <https://www.panorama.it/news/la-finestra-sul-cortile/venghino-signori-pensionati-venghino-al-sud/>

References

- Levelt, W. J. (2002). Phonological encoding in speech production: Comments on Jurafsky et al., Schiller et al., and van Heuven & Haan. *Laboratory Phonology*, 7.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38.
- Liberiamo. (2016, October 20). *I 10 errori grammaticali più comuni commessi dagli italiani*. <https://libreriamo.it/scuola/i-10-errori-grammaticali-piu-comuni-commessi-dagli-italiani/>
- Lieberman, E., Michel, J.-B., Jackson, J., Tang, T., & Nowak, M. A. (2007). Quantifying the evolutionary dynamics of language. *Nature*, 449(7163), 713–716.
- L'Impertinente, S. (2019, September 20). *Venghino Signori, venghino nella città dei “privati” e dei servizi pubblici spogliati*. <http://www.cn24tv.it/page/3362/venghino-signori-venghino-nella-citta-dei-privati-e-dei-servizi-pubblici-spogliati.html>
- Livius Patavinus, T. (1842). *Historiarum* (Vol. 2). Giuseppe Antonelli.

References

- Lloyd, P. M. (1987). *From Latin to Spanish: Historical phonology and morphology of the Spanish language* (Vol. 173). American Philosophical Society.
- Lo Duca, M. G. (2017, February 9). *Lettera di M.G. Lo Duca in risposta alla “proposta dei Seicento” sul declino della lingua italiana*.
<https://giscel.it/lettera-di-m-g-lo-duca-in-risposta-alla-proposta-dei-seicento-sul-declino-della-lingua-italiana/>
- Longrée, D. (2009). *The LASLA Database: Achievements and Projects*.
<http://hdl.handle.net/2268/122695>
- Lope Blanch, J. M. (1979). Investigaciones sobre dialectología mexicana. México, UNAM.
- Lorenz, D. (2013). *Contractions of English semi-modals: The emancipating effect of frequency*. Albert-Ludwigs-Universität Freiburg, Universitätsbibliothek.
- Lottiamo contro la scomparsa del congiuntivo*. (2007, November 26).
https://www.facebook.com/groups/6447658107/?tn-str=*F
- Lubbers Quesada, M. (1998). L2 acquisition of the Spanish subjunctive mood and prototype schema development. *Spanish Applied Linguistics*, 2(1), 1–23.

References

- Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics*, 39(3), 155–158.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19(1), 1.
- Lüdeling, A., & Kytö, M. (2008). *Corpus linguistics* (Vol. 1). Walter de Gruyter.
- Lynch, A. (1999). *The subjunctive in Miami Cuban Spanish: Bilingualism, contact, and language variability* [PhD Thesis]. University of Minnesota Minneapolis.
- Lyons, C. (1999). *Definiteness*. Cambridge University Press.
- Mackin, R., & Cowie, A. P. (1979). *Oxford Dictionary of Current Idiomatic English*. Oxford University Press.
- Magni, E. (2010). Mood and Modality. In P. Baldi & P. Cuzzolin (Eds.), *The foundations of Latin* (Vol. 117, pp. 193–275). Walter de Gruyter.
- Maiden, M. (2014). *Linguistic History of Italian*, A. Routledge.
- Malkiel, Y. (1992). Roger Wright (ed.). Latin and the Romance languages in the early Middle Ages. *Studies in Language*.

References

- International Journal Sponsored by the Foundation*
"Foundations of Language," 16(2), 448–467.
- Mańczak, W. (1957). Tendances générales des changements analogiques. *Lingua*, 7, 298–325.
- Manuwald, G. (2012). *Cicero, "Philippics" 3-9: Edited with Introduction, Translation and Commentary. Volume 1: Introduction, Text and Translation, References and Indexes. Volume 2: Commentary* (Vol. 30). Walter de Gruyter.
- Marcus, Gary F, Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29(3), 189–256.
- Marcus, Gary F, Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., & Clahsen, H. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, i–178.
- Marcus, Gary Fred. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science Learning, Development, and Conceptual Change*. MIT Press Cambridge, MA.

References

- Marquardt, D. (1970). *Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation, Technometrics*, 12, 591-612.
- Marslen-Wilson, W., & Tyler, L. K. (1997). Dissociating types of mental computation. *Nature*, 387(6633), 592.
- Marslen-Wilson, W., & Tyler, L. K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences*, 2(11), 428-435.
- Martineau, F. (1985). *Elision variable de (que) dans le parler d'Ottawa-Hull* [PhD Thesis]. University of Ottawa (Canada).
- Marslen, R. J., Theakston, A. L., Lieven, E. V., & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language, and Hearing Research*.
- Matras, Y. (2011). Grammaticalization and language contact. In H. Narrog & B. Heine (Eds.), *The Oxford handbook of grammaticalization*.

References

- Matsuda, K. (1998). On the conservatism of embedded clauses. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, 255–268.
- Matthews, P. H. (1974). *Morphology. An introduction to the theory of word structure*. Cambridge University Press.
- Mattino, I. (2017, October 20). *Gli strafalcioni: Da «qual'è» a «pò» alla «ceretta al linguine», la classifica degli errori più comuni*.
https://www.ilmattino.it/societa/persone/libreriamo_errori_grammaticali_italiani-3312216.html
- Mazzocchi, S. (2016, May 30). “Senza neanche un errore”: Come evitare strafalcioni e parlare sempre un italiano corretto.
https://www.repubblica.it/rubriche/passaparola/2016/05/30/news/senza_neanche_un_errore-140938963/
- McCarthy, J. J. (2002). *A thematic guide to Optimality Theory*. Cambridge University Press.
- McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, 6(11), 465–472.

References

- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375.
- McCormick, S. F., Brysbaert, M., & Rastle, K. (2009). Is morphological decomposition limited to low-frequency words? *The Quarterly Journal of Experimental Psychology*, 62(9), 1706–1715.
- McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, 44(3), 295–322.
- Meillet, A. (1912). *L'évolution des formes grammaticales*.
- Messaggero, I. (2017, October 19). *Gli strafalcioni: Da “qual’è” a “pò” alla “ceretta al linguine”, la classifica degli errori più comuni*.
https://www.ilmessaggero.it/primopiano/cronaca/libreriamo_errori_grammaticali_italiani-3312167.html
- Meunier, F., & Segui, J. (1999). Morphological priming effect: The role of surface frequency. *Brain and Language*, 68(1–2), 54–60.

References

- Mikheev, A. (1997). Automatic rule induction for unknown-word guessing. *Computational Linguistics*, 23(3), 405–423.
- Milin, P., \DJur\djević, D. F., & del Prado Martín, F. M. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, 60(1), 50–64.
- Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, H. (2017). Discrimination in lexical decision. *PloS One*.
- Milin, P., Kuperman, V., Kostic, A., & Baayen, R. H. (2009). Paradigms bit by bit: An information theoretic approach to the processing of paradigmatic structure in inflection and derivation. *Analogy in Grammar: Form and Acquisition*, 214–252.
- Milin, P., Smolka, E., & Feldman, L. B. (2018). *Models of lexical access and morphological processing*.
- Milroy, J., & Milroy, L. (1978). Belfast: Change and variation in an urban vernacular. *Sociolinguistic Patterns in British English*, 19, 19–36.

References

- Moneglia, M. (2003). L'italiano come risorsa romanza nel corpus multilingue C-ORAL-ROM. *Atti Del Convegno Di Studi 'Il Parlato Italiano.'*
- Moneglia, M. (2005). I corpora dell'italiano parlato di LABLITA: criteri di costituzione, unità di analisi e comparabilità dei dati linguistici orali. *Burr, E.(a Cura Di)*, 213–231.
- Montermini, F. (2001). The unitary base hypothesis and the semantics of word-formation rules. *First International Workshop on Generative Approaches to The*, 26–28.
- Montrul, S. (2007). Interpreting mood distinctions in Spanish as a heritage language. In K. Potowski & R. Cameron (Eds.), *Spanish in contact policy, Social and linguistic inquiries*. John Benjamins.
- Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *International Journal of Bilingualism*, 13(2), 239–269.
- Moscoso del Prado Martin, F., & Baayen, R. H. (2005). *Breaking the tyranny of learning: A broad-coverage distributed connectionist model of visual word recognition*.

References

- Mousikou, P., & Rastle, K. (2015). Lexical frequency effects on articulation: A comparison of picture naming and reading aloud. *Frontiers in Psychology*, 6, 1571.
- Napoli, D. J., & Vogel, I. (1990). The conjugations of Italian. *Italica*, 67(4), 479–502.
- New, B., Brysbaert, M., Segui, J., Ferrand, L., & Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language*, 51(4), 568–585.
- Niedzielski, N. A., & Preston, D. R. (2010). *Folk linguistics* (Vol. 122). Walter de Gruyter.
- Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language*, 14(2), 177–202.
- Ocampo, F. (1990). El subjuntivo en tres generaciones de hablantes bilingües. *Spanish in the United States: Sociolinguistic Issues*, 39–48.
- O'Donnell, T. J. (2015). *Productivity and reuse in language: A theory of linguistic computation and storage*. MIT Press.

References

- Olivetti, E. (Ed.). (2003). *Dizionario Italiano Olivetti [Electronic Resource]*. <https://www.dizionario-italiano.it/>
- Orsolini, M., Fanari, R., & Bowles, H. (1998). Acquiring regular and irregular inflection in a language with verb classes. *Language and Cognitive Processes*, 13(4), 425–464.
- Orsolini, M., & Marslen-Wilson, W. (1997). Universals in morphological representation: Evidence from Italian. *Language and Cognitive Processes*, 12(1), 1–47.
- Padula, C. (2016, July 15). *I 10 errori grammaticali più comuni: Gli “strafalcioni” più frequenti commessi nella lingua italiana*. <https://www.nanopress.it/cultura/2016/07/15/10-errori-grammaticali-piu-comuni/138493/>
- Pagliara, F. (2018). L’analisi dei bisogni: Uno studio sull’italiano come lingua della comunicazione accademica. *IL TASK NELL’INSEGNAMENTO DELLE LINGUE. PERCORSI TRA RICERCA E DIDATTICA AL CLA DI ROMA TRE*, 77–104.
- Pallotta, A. (2006). *Mostruosamente...Fantozzi*. Un mondo a parte.
- Palmer, F. R. (1986). *Mood and Modality*. Cambridge University Press.

References

- Palumbo, V. (2016, September 9). *La lotta con il congiuntivo: Di Maio soltanto l'ultimo della catena*.
https://www.corriere.it/politica/16_settembre_09/lotta-il-congiuntivo-maio-soltanto-l-ultimo-catena-fcbc0c20-7666-11e6-b673-b2cde5239b18.shtml
- Paradis, J., Crago, M., & Genesee, F. (2006). Domain-general versus domain-specific accounts of specific language impairment: Evidence from bilingual children's acquisition of object pronouns. *Language Acquisition*, 13(1), 33–62.
- Patota, G. (2020, March 23). *Interrogative Dirette*.
[http://www.treccani.it/enciclopedia/interrogative-dirette_\(Enciclopedia-dell'Italiano\)/](http://www.treccani.it/enciclopedia/interrogative-dirette_(Enciclopedia-dell'Italiano)/)
- Patota, G., & Della Valle, V. (2016). *Senza neanche un errore*. Sperling & Kupfer.
- Pearson, K. (1989). Chi-square test (x2). *Phil. Mag. Ser.*, 5, 50–157.
- Peirce, J. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2:

References

- Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203.
- Pfänder, S., & Behrens, H. (n.d.). Experience counts: An introduction to frequency effects in. In Behrens (Ed.), *Experience counts: Frequency effects in language*.
- Picallo, C. (1985). Opaque Domains, PhD diss. *City University of New York*.
- Picallo, M. C. (1984). The Infl node and the null subject parameter. *Linguistic Inquiry*, 75–102.
- Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology* (Vol. 7, pp. 101–139). Mouton.
- Pietrandrea, P. (2007). The grammatical nature of some epistemic-evidential adverbs in spoken Italian. *Italian Journal of Linguistics*, 19(1), 39–64.
- Pietrandrea, P. (2012). The conceptual structure of irreality: A focus on non-exclusion-of-factuality as a conceptual and a linguistic category. *Language Sciences*, 34(2), 184–199.

References

- Pietropaolo, C. (2014). *Evolution of the subjunctive mood in Italian: Morpho-syntactic features predicting mood selection* [Master's Thesis]. Albert-Ludwigs-Universität at Freiburg.
- Pinker, S. (1991). Rules of Language. *Science*, 253, 530–536.
- Pinker, S. (1995a). Language acquisition. *Language: An Invitation to Cognitive Science*, 1, 135–182.
- Pinker, S. (1995b). Why the child holded the baby rabbits: A case study in language acquisition. *An Invitation to Cognitive Science*, 1, 107–133.
- Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28(1–2), 73–193.
- Pinker, S., & Prince, A. (1994a). Regular and irregular morphology and the psychological status of rules of grammar. *The Reality of Linguistic Rules*, 321, 51.

References

- Pinker, S., & Prince, A. (1994b). Regular and irregular morphology and the psychological status of rules of grammar. *The Reality of Linguistic Rules*, 321, 51.
- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6(11), 456–463.
- Pinkster, H. (1986). Three notes on the Latin subjunctive'. *Revue Informatique et Statistique Dans Les Sciences Humaines*, 22, 147–156.
- Pinkster, H. (1990). *Latin syntax and semantics*. Taylor & Francis.
- Pinkster, H. (2015). *The Oxford Latin Syntax: Volume 1: The Simple Clause*. OUP Oxford.
- Pizzuto, E., Caselli, M. C., & Levy, Y. (1994). The acquisition of Italian verb morphology in a cross-linguistic perspective. *Other Children, Other Languages: Issues in the Theory of Language Acquisition*, 137–187.
- Plautus, T. M. (1847). *Le commedie di M. Accio Plauto con traduzione e note*. G. Antonelli.
- Poletto, C. (2000). *The higher functional field: Evidence from Northern Italian dialects*. Oxford University Press.

References

- Poletto, C. (2001). Complementizer deletion and verb movement in standard Italian. *G. Cinque e GP Salvi, Current Studies in Italian Syntax, Amsterdam, Elsevier*, 265–286.
- Poplack, S. (1989). The care and handling of a megacorporus: The Ottawa-Hull French project. *Language Change and Variation*, 4.
- Poplack, S. (1990). Prescription, intuition et usage: Le subjonctif français et la variabilité inhérente. *Langage & Société*, 54(1), 5–33.
- Poplack, S. (1991). The inherent variability of the French subjunctive. *Current Issues in Linguistic Theory*, 235–263.
- Poplack, S. (1995). The sociolinguistic dynamics of apparent convergence. In G. Guy, J. Baugh, & D. Schiffrin (Eds.), *Towards a social science of language: Papers in honor of William Labov* (pp. 285–309). Amsterdam: John Benjamins.
- Poplack, S., & Bourdages, J. S. (2005). Corpus du français en contexte: Milieux scolaire et social. *University of Ottawa. SSHRC Research Grant*, 410–2005.

References

- Poplack, S., Cacoullos, R. T., Dion, N., de Andrade Berlinck, R., Digesto, S., Lacasse, D., & Steuck, J. (2018). Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. *Manuals in Linguistics: Romance Sociolinguistics*. Berlin: De Gruyter, 217–252.
- Poplack, S., Lealess, A., & Dion, N. (2013). The evolving grammar of the French subjunctive. *International Journal of Latin and Romance Linguistics*, 25(1), 139–195.
- Poplack, S., & St-Amand, A. (2007). A real-time window on 19th-century vernacular French: The Récits du français québécois d'autrefois. *Language in Society*, 36(5), 707–734.
- Portner, P. (2004). The semantics of imperatives within a theory of clause types. *Semantics and Linguistic Theory*, 14, 235–252.
- Porto Dapena, J. A. (1991). *Del indicativo al subjuntivo: Valores y usos de los modos del verbo*. Arco Libros.
- Prandi, M. (2002). C'è un valore per il congiuntivo? *C'è Un Valore per Il Congiuntivo?*, 1000–1016.

References

- Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8(1), 1–56.
- Price, G. (1998). The French Language: Present and past. Rev. Ed. *London: Grant & Cutler*.
- Progovac, L. (1993). Subjunctive: The (mis) behavior of anaphora and negative polarity. *The Linguistic Review*, 10(1), 37–60.
- Progovac, L. (1994). Positive and negative polarity: A binding approach. *Cambridge Studies in Linguistics*, 38.
- Quer, J. (1997). In the cause of subjunctive. *Linguistics in the Netherlands*, 14(1), 171–182.
- Quer, J. (1998). *Mood at the Interface* [PhD Thesis]. Doctoral dissertation, Netherlands Graduate School of Linguistics, UiL OTS.
- Quer, J. (2009). Twists of mood: The distribution and interpretation of indicative and subjunctive. *Lingua*.
- Quotidiano, I. F. (2018, November 24). “Non so perché i giornali scrivino...”, *Conte alle prese con la ‘congiuntivite’*. *E lo strafalcione fa il giro dei social*.

References

- <https://www.ilfattoquotidiano.it/2018/11/24/non-so-perche-i-giornali-scrivino-conte-alle-prese-con-la-congiuntivite-e-lo-stra-falcione-fa-il-giro-dei-social/4789523/>
- Rácz, P. (2013). *Salience in sociolinguistics: A quantitative approach* (Vol. 84). Walter de Gruyter.
- Ramscar, M., & Baayen, H. (2013). Production, comprehension, and synthesis: A communicative perspective on language. *Frontiers in Psychology*, 4, 233.
- Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends in Cognitive Sciences*, 11(7), 274–279.
- Raveh, M. (2002). The contribution of frequency and semantic similarity to morphological processing. *Brain and Language*, 81(1–3), 312–325.
- Renzi, L., Salvi, G., & Cardinaletti, A. (1991). *Grande grammatica italiana di consultazione* (Vol. 2). Il mulino.
- Repubblica, L. (2017, January 14). *M5s, Di Maio e il congiuntivo: Tre tentativi per un tweet*.
https://www.repubblica.it/politica/2017/01/14/foto/m5s_di_m

References

- aio_sbaglia_tre_volte_il_congiuntivo_l_ironia_su_twitter-156007460/1/#1
- Rissa, A. (2015). *Il culo non esiste solo per andare di corpo*. Antologia della letteratura greca e latina, Il Melangolo.
- Rivero, M.-L. (1980). On left-dislocation and topicalization in Spanish. *Linguistic Inquiry*, 11(2), 363–393.
- Rizzi, L. (1990). On the anaphor-agreement effect. *Rivista Di Linguistica*, 2(1), 27–42.
- Rizzo, S. (2016). *La repubblica dei brocchi: Il declino della classe dirigente italiana*. Feltrinelli.
- Rosemeyer, M. (2015). How usage rescues the system: Persistence as conservation. In A. Adli, M. García, & G. Kaufmann (Eds.), *Storage and computation in the language faculty*. Berlin, New York: De Gruyter.
- Rosemeyer, M. (2016). Modeling frequency effects in language change. *Experience Counts: Frequency Effects in Language*, 54, 175.

References

- Rosemeyer, M., & Schwenter, S. A. (2019). Entrenchment and persistence in language change: The Spanish past subjunctive. *Corpus Linguistics and Linguistic Theory*, 15(1), 167–204.
- Ross Veidmark, R., & Umaña Aguiar, J. (1991). La desaparición del subjuntivo español y sus implicaciones para el cambio lingüístico. *Revista de Filología y Lingüística de La Universidad de Costa Rica*, 17(1–2), 193–202.
- Rossi, F. (2002). *La lingua in gioco: Da Totò a lezione di retorica* (Vol. 3). Bulzoni.
- Roussou, A. (2000). On the left periphery. *Journal of Greek Linguistics*, 1, 63–92.
- Rumelhart, D. E., & McClelland, J. L. (1986). *On learning the past tenses of English verbs*.
- Ruwet, N. (1984). Je veux partir/* Je veux que je parte. A propos de la distribution des complétives à temps fini et des compléments à l’infinitif en français. *Cahiers de Grammaire*, 7, 75–138.
- Ruwet, N. (1991). *Syntax and human experience*. University of Chicago Press.

References

- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical cues in language acquisition: Word segmentation by infants. *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, 376–380.
- Sankoff, D., & Laberge, S. (1978). The linguistic market and the statistical explanation of variability. *Linguistic Variation: Models and Methods*, 239–250.
- scuola, N. della. (2017, February 10). *Una risposta alla “proposta dei Seicento” sul declino della lingua italiana*.
<https://www.notiziedellascuola.it/news/2017/febbraio/una-risposta-alla-proposta-dei-seicento-sul-declino-della-lingua-italiana>
- Scuola, R. (2019, January 28). *I rilievi della Crusca. Ceretta al linguine e cortelli affilati, ecco gli errori di grammatica più comuni*.
<https://www.corriere.it/scuola/universita/cards/ceretta-linguine-cortelli-affilatigli-errori-grammatica-piu-comuni/congiuntivo-questo-sconosciuto.shtml>
- Sera, C. della. (2014, March 12). *Renzi-venditore: «Venghino, signori venghino»*. <https://video.corriere.it/auto-blu-renzi-fa>

References

- venditore-venghino-signori-venghino/0ac7c352-aa10-11e3-9476-764b3ca84ea2
- Sera, I. C. della. (2018, June 5). *Il congiuntivo sbagliato di Conte: «Si arricchischino». Ma poi si corregge subito.*
<https://video.corriere.it/congiuntivo-sbagliato-conte-si-aricchischino-ma-poi-si-corregge-subito/23449f6c-68df-11e8-861b-78f7b945ff41>
- Talmy, G. (1971). Historical syntax and synchronic morphology: An archaeologist's field trip. *Seventh Regional Meeting, Chicago Linguistic Society*, 394–415.

German summary

Die Dissertation analysiert die Verwendung und Produktivität des italienischen Subjunktivs. Darüber hinaus wird der Subjunktiv als Fallstudie untersucht, um herauszufinden wie produktive morphologische Regeln entstehen, und wie sie aufrechterhalten werden angesichts morphologischer Variation und Vereinfachung. Das Buch integriert sehr unterschiedliche Bereiche der Sprachforschung, wobei widersprüchliche Ansichten überprüft und kritisch bewertet werden, um eine umfassende Analyse des Subjunktivs zu präsentieren, die bisher nie bereitgestellt wurde.

Das Konzept der Modus-Variation wird vorgestellt. Modus-Variation bezeichnet die Verwendung verschiedener Modalformen, insbesondere die Verwendung des Subjunktivs, des Indikativs und des Konditionals in Nebensätzen, die gemäß der Sprachnorm den Subjunktiv verlangen würden (*Locus* der Variation). Hierbei handelt es sich um ein Phänomen der gesprochenen Sprache, obwohl es heutzutage auch in der informellen schriftlichen Kommunikation üblich ist.

Dieses Phänomen der morphologischen Variation wirft folgende Fragen zur internen Grammatik des Subjunktivs der Sprecher auf:

1. Wie kann die Verwendung des Subjunktivs beschrieben und verstanden werden?
2. Ist die morphologische Variation zufällig oder reagiert sie auf sprachinterne und kognitive Prinzipien?

Ziel der Dissertation ist es, solche Prinzipien zu identifizieren.

Insbesondere wird Folgendes untersucht:

- a. Beeinflussen semantische Faktoren die Subjunktiv- / Indikativverteilung in der Sprache?
- b. Wie beeinflusst das Hauptverb (Governor) die Wahrscheinlichkeit, dass der Subjunktiv im Nebensatz auftritt?
- c. Ist der Subjunktiv in bestimmten Phrasen verankert?

Diese Forschungsfragen wurden mittels einer Korpusstudie untersucht. Die Studie ergab, dass kein einzelnes Hauptverb (Governor) ausschließlich den Subjunktiv verlangt, nicht einmal die desiderativen/volitiven Verben, die nach einigen Theorien die zentrale Subjunktivbedeutung darstellen (siehe Quer, 1997, 2009; Kempchinsky, 2009). Am interessantesten ist, dass Governor-Verben, die derselben semantischen Klasse angehören oder Quasi-Synonyme sind, unterschiedliche Korrelationen mit dem Subjunktiv aufweisen (siehe zum Beispiel die epistemischen Verben *credere* „glauben“, *pensare* „denken“).

Dieses Ergebnis widerspricht Theorien, die behaupten, dass die Modus-Variation hauptsächlich semantisch motiviert ist (siehe Picallo, 1984; Farkas, 1992a, 1992b; Quer, 1997, 2009; Giorgi and Pianesi, 1997, 2002, 2004a, 2004b u.a.). Darüber hinaus ergab sich aus der Korpusstudie, dass heutzutage nur zwei Hauptverben (Governors) den Subjunktiv konsequent auswählen, nämlich die Verben *volere* „wollen“ und *bisognare* „notwendig sein“ in der Form *bisogna* „es ist

notwendig“. Ähnliche Ergebnisse wurden für Französisch berichtet (siehe die Korpusstudien in Poplack, 1991; Poplack et al., 2013). Dieses Ergebnis stützt die Hypothese, dass der Subjunktiv in formelhaften Phrasen oder „Schemata“ (siehe Bybee & Thompson, 1997) verankert ist („Lexikalisierung“).

Lexikalisierung bezeichnet einen Prozess, der in der historischen Linguistik als Teil der Grammatikalisierung verstanden wird (Lehmann, 2015, p. 149): Lexikalische Elemente, die häufig zusammen auftreten, werden routiniert. Gleichzeitig verlieren die einzelnen Bestandteile der Phrase ihre unabhängige Bedeutung. Lexikalisierung ist ein Indikator für eine Abnahme an Produktivität des morphologischen Systems. Kristallisierte Phrasen stellen sicher, dass der Subjunktiv als veraltete grammatikalische Form weiterhin verwendet wird. Sprecher verwenden den Modus jedoch nicht mehr frei, um neue Bedeutungen zu schaffen.

Letztendlich wurde der Subjunktiv als Fallstudie verwendet, um zu untersuchen, wie Frequenz das morphologische Lernen und die Produktivität morphologischer Regeln beeinflusst. Insbesondere wurden folgende Fragen untersucht:

- d. Wie verallgemeinern Sprecher morphologische Muster auf neue Wörter?
- e. Wird die Verallgemeinerung sowohl regelmäßiger als auch unregelmäßiger morphologischer Muster durch die Ähnlichkeit mit vorhandenen Wörtern beeinflusst?

- f. Beeinflusst die Häufigkeit des Auftretens der morphologischen Muster ihre Produktivität?

Zwei wichtige Modelle des morphologischen Lernens machen unterschiedliche Vorhersagen. Das „Dual-Mechanism“ Modell (Pinker & Prince, 1988; Pinker, 1991) geht davon aus, dass nur unregelmäßige Prozesse in Analogie zu vorhandenen Wörtern verallgemeinert werden und somit empfindlich auf die phonologischen Eigenschaften und die Statistik des Lexikons reagieren. Regelmäßige morphologische Prozesse werden mittels einer kontextunabhängigen Regel angewendet, die gegenüber den Eigenschaften des Lexikons unempfindlich ist.

Analogiebasierte Modelle (*Analogy-based models*) gehen dagegen davon aus, dass die Erfahrung mit der Sprache, die das Kind und der erwachsene Sprecher sammeln, für die Gestaltung ihres morphologischen Wissens von grundlegender Bedeutung ist (Rumelhart & McClelland, 1986, McClelland & Patterson, 2002; Bybee, 1995). Insbesondere nach diesem Modell werden alle morphologischen Regeln durch Extrahieren von Ähnlichkeiten zwischen vorhandenen Wörtern gebildet. Die Häufigkeit, mit der eine bestimmte morphologische Änderung in der Sprache auftritt, ist von grundlegender Bedeutung für die Bestimmung, ob die Regel produktiv ist und daher auch auf neue Wörter angewendet wird.

Zwei Analogie-basierte Modelle unterscheiden sich hinsichtlich der Frequenzmaße, welche die Produktivität morphologischer Muster direkt beeinflusst. Nach dem Netzwerkmodell (*Network model*, [Bybee,

1995]) beeinflusst nur die Anzahl der verschiedenen Wörter, die ein morphologisches Muster annehmen, direkt die Produktivität. Dieses Maß wird als Typfrequenz bezeichnet. Die Hypothese basiert auf der Beobachtung, dass hochfrequente Wörter die Produktivität unregelmäßiger Muster nicht zu verbessern scheinen (z. B. unregelmäßige Vergangenheitsformen in Englisch [Bybee, 1995, p. 428]). Nur wenn das morphologische Schema eine große Anzahl von Wörtern umfasst, wird es verallgemeinert.

Konnektionistische Modelle (Rumelhart & McClelland, 1986, McClelland & Patterson, 2002) dagegen sagen voraus, dass die absolute Häufigkeit, mit der ein morphologischer Wechsel auftritt, die Wahrscheinlichkeit beeinflusst, dass sie angewendet wird, um neue Wortformen abzuleiten. Dies bedeutet, dass sowohl die Anzahl der verschiedenen lexikalischen Elemente (Typfrequenz) als auch die absolute Häufigkeit (Tokenfrequenz) eine Rolle für die Produktivität spielen.

Diese Hypothesen wurden unter Verwendung einer mündlichen Produktionsaufgabe getestet (Wug-Experiment), bei der erwachsene italienische MuttersprachlerInnen aufgefordert wurden, Subjunktive und Infinitive neuer Verben abzuleiten. Diese zwei Modi wurden ausgewählt, weil sie sich in entscheidenden Aspekten hinsichtlich der Häufigkeit unterscheiden, mit der sie in der Sprache erlebt werden. Die Mehrzahl der Subjunktive tritt bei hochfrequenten Verben auf, die hauptsächlich dem unregelmäßigen morphologischen Muster (zweite, dritte und vierte Konjugation) folgen. Somit wäre das unregelmäßige

Muster, das ausschließlich auf der Token-Frequenz basiert, das häufigste in der Sprache. Das regelmäßige Subjunktivmuster (erste Konjugation) hat die höchste Typfrequenz. Im Gegensatz dazu weist das Infinitivmuster der ersten Konjugation die höchste Typ- und die höchste Tokenfrequenz auf.

Die Ergebnisse des Experiments bestätigen die Vorhersagen Analogie-basierter Modelle. Es wurde festgestellt, dass die Anwendung aller morphologischen Prozesse durch die phonologische Ähnlichkeit der neuen Verben mit vorhandenen Wörtern beeinflusst wird. Dies war sowohl für die Subjunktive als auch für die Infinitive der Fall. Es gab keinen Unterschied in der Art und Weise, wie reguläre und unregelmäßige Prozesse angewendet wurden, ein Ergebnis, welches das „Dual-Mechanism“-Modell untergräbt.

Zusätzlich wurde festgestellt, dass ein kombiniertes Maß von Typ- und Tokenfrequenz die Produktivität morphologischer Muster beeinflusst. Die Typhäufigkeit allein überschätzte die Produktion der regelmäßigen Infinitive und Subjunktive der ersten Konjugation. Im Fall des Subjunktivs war die Überschätzung beträchtlich: Die Typhäufigkeit war bei weitem das schlechteste Maß und schnitt viel schlechter ab als die Tokenfrequenz. Gleiches gilt für die Anwendung der unregelmäßigen Muster: Die Typhäufigkeit hat die Erzeugung dieser Formen unterschätzt, insbesondere im Fall des Subjunktivs. Dieses Ergebnis bestätigt die Vorhersagen der konnektionistischen Modelle (Rumelhart & McClelland, 1986) und widerspricht der Hypothese des Netzwerkmodells (Bybee, 1995).