



UNIVERSITÀ  
di **VERONA**



**SORBONNE  
UNIVERSITÉ**  
CRÉATEURS DE FUTURS  
DEPUIS 1257

UNIVERSITÀ DEGLI STUDI DI VERONA  
SORBONNE UNIVERSITÉ

*Dottorato di Ricerca in Letterature Straniere, Lingue e Linguistica*  
*Doctorat en linguistique (École Doctorale 433 Concepts et Langues)*  
*Cycle: XXXVI*

**Auxiliary Selection in Italian and French**  
*A comparative study of the so-called peripheral verbs*

By Greta Viale

DOCTORAL THESIS IN LINGUISTICS

Supervised by Stefan Rabanus and Anne Carlier

Publicly defended at the University of Verona on February 28<sup>th</sup>, 2025

Examination committee composed of:

Professor Stefan RABANUS (Università degli Studi di Verona), Thesis Supervisor

Professor Anne CARLIER (Sorbonne Université), Thesis Supervisor

Professor Lea SCHÄFER (Universität Kassel), Reviewer

Professor Mario SQUARTINI (Università degli Studi di Torino), Reviewer

Professor Achim STEIN (Universität Stuttgart), Examiner







## ABSTRACT

This thesis investigates the intricate phenomenon of auxiliary selection in two Romance languages, namely Italian and French. The primary objective is to elucidate the factors that influence the choice between the auxiliaries ‘be’ and ‘have’ in the formation of the perfect tense. The study focuses on verbs that can select both auxiliaries, commonly known as peripheral verbs (Sorace 2000), which, despite extensive individual examination, have not been comprehensively analyzed (Giancarli 2015).

The central research questions addressed are: What characteristics enable these verbs to select both auxiliaries? Which factors determine the predominance of one auxiliary over the other? What is the relative weight of factors such as agentivity and telicity in auxiliary selection (Sorace 2000)?

For the first time, this research systematically explores auxiliary selection in Italian and French using corpus analysis and natural language processing (NLP). By integrating these methods, the study aims to identify the most significant factors influencing auxiliary choice in intransitive verbs with auxiliary alternation. The research combines qualitative analysis of manually annotated occurrences from SketchEngine (Kilgarriff et al. 2014) with quantitative analysis using statistical models to determine the most significant parameters in auxiliary selection.

The findings reveal the paramount importance of semantic, syntactic, and morphological aspects in the choice of ‘be’ or ‘have’. Notably, telicity is found to be less relevant for these verbs. The study also highlights significant differences between Italian and French. Italian verbs are categorized into full verbs and semi-auxiliaries. For full verbs, internal cause and human traits are crucial factors in auxiliary selection. For semi-auxiliary verbs, the type of infinitive and the human trait associated with particular infinitives are shown to be significant. In French, the type of construction heavily influences auxiliary choice.

By providing comprehensive answers to previously unexplored areas, this study aligns with and extends the existing literature. It significantly enhances our understanding of verb categorization and auxiliary selection, with substantial implications for both theoretical and applied linguistics. Furthermore, it underscores the importance of integrative methodological approaches for analyzing complex linguistic phenomena.

Key words: Auxiliary Selection, Peripheral Verbs, Internal Cause, Corpus Analysis, NLP



# CONTENTS

<b>LIST OF ABBREVIATIONS .....</b>	<b>9</b>
<b>INTRODUCTION.....</b>	<b>11</b>
<b>CHAPTER 1: PREVIOUS SYNTACTIC AND SEMANTIC APPROACHES.....</b>	<b>17</b>
1.1 THE UNACCUSATIVE HYPOTHESIS: THE SYNTACTIC APPROACH (PERLMUTTER, BURZIO).....	17
1.2 UNACCUSATIVE VS. UNERGATIVE VERBS: AT THE INTERFACE BETWEEN SYNTAX AND SEMANTICS (GRIMSHAW, LEVIN & RAPPAPORT) .....	18
1.3 UNACCUSATIVE VS. UNERGATIVE VERBS: A SEMANTIC APPROACH (VAN VALIN).....	22
1.4 AUXILIARY SELECTION HYPOTHESIS (SORACE) .....	30
1.5 A BINARY FEATURE APPROACH OF THE ASH (BENTLEY & EYTHÓRSSON.) .....	36
1.6 LIMITS OF THE ASH APPROACH ACCORDING TO GIANCARLI (2015).....	39
1.7 CONCLUSION: SYNTHESIS AND PATHWAY TO METHODOLOGY .....	44
<b>CHAPTER 2: METHODOLOGY.....</b>	<b>47</b>
2.1 CORPUS DATA.....	48
2.1.1 CHOICE OF THE CORPORA .....	48
2.1.2 METHODS FOR QUALITATIVE CORPUS ANALYSIS .....	49
2.1.3 PARAMETERS FOR QUALITATIVE CORPUS ANALYSIS.....	53
2.2 FROM QUALITATIVE TO QUANTITATIVE CORPUS ANALYSIS: STATISTICAL METHODS .....	62
2.3 THE FRESH INSIGHTS FROM THIS ANALYSIS.....	63
<b>CHAPTER 3: CORPUS ANALYSIS OF ITALIAN VERBS .....</b>	<b>65</b>
3.1. CHANGE-OF-STATE: <i>CAMBIARE</i> .....	68
3.2 CHANGE OF POSSESSION: <i>CEDERE</i> .....	79
3.3 VERBS OF SOUND EMISSION: <i>SUONARE</i> .....	84
3.4 EMOTIONAL STATE VERBS.....	90
3.4.1 <i>Contare</i> .....	90
3.4.2 <i>Pesare</i> .....	97
3.5 VERBS OF CONTIGUOUS LOCATION : <i>PREVALERE</i> .....	106
3.6 VERBS OF FAILURE: <i>FALLIRE</i> .....	111
3.7 VERBS INDICATING A STAGE OF A PROCESS.....	117
3.7.1 VERBS INDICATING THE START OF A PROCESS: <i>iniziare</i> .....	117
3.7.2 VERBS INDICATING THE START OF A PROCESS: <i>cominciare</i> .....	123
3.7.3 VERBS INDICATING THE PROGRESSION OF A PROCESS: <i>procedere</i> .....	126
3.7.4 VERBS INDICATING THE CONTINUATION OF AN ACTION: <i>continuare</i> .....	133
3.7.5 VERBS INDICATING THE CONTINUATION OF A PROCESS: <i>proseguire</i> .....	139
3.7.6 VERBS INDICATING THE COMPLETION OF A PROCESS: <i>finire</i> .....	144
3.8 INTERMEDIARY RESULTS .....	149
3.9 ASPECTUAL VERBS WITH INFINITIVE.....	151
3.9.1 <i>Iniziare</i> .....	154
3.9.2 <i>Cominciare</i> .....	161
3.9.3 <i>Continuare</i> .....	169
3.9.4 <i>Finire</i> .....	177
3.10 INTERMEDIARY RESULTS: THE SEMI-AUXILIARIES OF ASPECT IN INFINITIVE .....	185
3.11 CONCLUDING REMARKS ON AUXILIARY SELECTION IN ITALIAN.....	187
<b>CHAPTER 4: CORPUS ANALYSIS OF FRENCH VERBS .....</b>	<b>191</b>
4.1 SPATIAL MOTION VERBS .....	195
4.1.1 <i>Passer</i> .....	195
4.1.2 <i>Monter</i> .....	206
4.1.3 <i>Baisser</i> .....	222
4.1.4 <i>Descendre</i> .....	229
4.1.5 <i>Conclusion: Spatial Motion verbs</i> .....	245

4.2 CHANGE-OF-STATE VERBS .....	246
4.2.1 <i>Changer</i> .....	246
4.2.2 <i>Mûrir</i> .....	254
4.2.3 <i>Diminuer</i> .....	261
4.2.4 <i>Exploser</i> .....	268
4.2.5 <i>Conclusion: Change-of-State verbs</i> .....	276
4.3 VERBS OF APPEARANCE .....	277
4.3.1 <i>Paraître</i> .....	277
4.3.2 <i>Apparaître</i> .....	282
4.3.3 <i>Conclusion: Verbs of Appearance</i> .....	290
4.4 CONCLUDING REMARKS ON AUXILIARY SELECTION IN FRENCH.....	291
<b>CHAPTER 5: STATISTICAL ANALYSIS.....</b>	<b>295</b>
5.1 INTRODUCTION: CHAID .....	295
5.2 INTRODUCTION: RANDOM FOREST.....	297
5.3 CHAID AND RANDOM FOREST FOR ITALIAN VERBS .....	297
5.4 CHAID AND RANDOM FOREST FOR ITALIAN ASPECTUAL VERBS .....	303
5.5 CHAID AND RANDOM FOREST FOR FRENCH VERBS .....	306
5.5.1 <i>CHAID AND RANDOM FOREST WITH SPATIAL MOTION VERBS (passer, monter, baisser, descendre, diminuer)</i> .....	309
5.5.2 <i>CHAID AND RANDOM FOREST WITH CHANGE-OF-STATE VERBS (changer, mûrir, exploser)</i> .....	313
5.5.3 <i>CHAID AND RANDOM FOREST WITH VERBS OF APPEARANCE (paraître, apparaître)</i> .....	316
5.6 CONCLUDING REMARKS .....	318
<b>CHAPTER 6: COMPARATIVE ANALYSIS OF AUXILIARY SELECTION IN ITALIAN AND FRENCH.....</b>	<b>321</b>
6.1 AUXILIARY DISTRIBUTION IN ITALIAN: THE OUTPUTS OF CORPUS ANALYSIS AND STATISTICAL ANALYSIS.....	321
6.2 AUXILIARY DISTRIBUTION IN FRENCH: THE OUTPUTS OF CORPUS ANALYSIS AND STATISTICAL ANALYSIS .....	322
6.3 A COMPARATIVE ANALYSIS OF AUXILIARY DISTRIBUTION IN ITALIAN AND FRENCH .....	324
<b>CONCLUSION .....</b>	<b>327</b>
<b>REFERENCES.....</b>	<b>337</b>

## LIST OF ABBREVIATIONS

1 - First person  
2 - Second person  
3 - Third person  
ADJ - Adjective  
ADV - Adverb  
ART - Article  
COND - Conditional  
DAT - Dative  
DEF - Definite  
DET - Determiner  
F – Feminine  
FUT – Future  
INF - Infinitive  
INDF - Indefinite  
IPFV - Imperfective  
M - Masculine  
PL - Plural  
POSS - Possessive  
PRF - Perfect  
PRS - Present  
PTCP – Participle  
REFL - Reflexive  
SG - Singular  
SUBJ - Subjunctive



## INTRODUCTION

In Italian and French, the perfect tense can be constructed using either the auxiliary ‘be’ or ‘have’. The study of auxiliary selection in these two Romance languages provides fascinating insights into the intricate relationship between syntax, lexical semantics, and morphology. Understanding the factors that influence this choice offers valuable contributions to linguistic theory.

Auxiliary selection is not a random phenomenon; rather, it is governed by specific syntactic and semantic principles. In both Italian and French, transitive verbs consistently form the perfect tense with the auxiliary ‘have’:

- (1) a. ITALIAN: *Giovanni ha attraversato la barriera*  
b. FRENCH : *Jean a traversé la barrière*  
‘John crossed the fence.’

However, intransitive verbs display auxiliary split. These verbs can select either ‘be’ or ‘have’ as their auxiliary, and the choice often depends on the inherent properties of the verb itself. The present study focuses on the differences in auxiliary selection between Italian and French. An interesting example is the verb ‘blush’ selecting ‘be’ in Italian (*Giorgio è arrossito* ‘George blushed’), whereas French prototypically requires ‘have’:

- (2) a. ITALIAN: *Giorgio è arrossito*  
b. FRENCH: *Georges a rougi*  
‘George blushed’

Another topic of interest are verbs exhibiting auxiliary alternation within a single language, illustrated by the Italian verb *saltare* ‘jump’ in (3):

(3) ITALIAN:

a. *Ida ha saltato sul letto*

b. *Ida è saltata sul letto*

‘Ida jumped on the bed.’

This variation in auxiliary selection—both across languages and within the same language—raises several important questions. What are the underlying reasons for this split? Are they syntactic or semantic? Why do some languages show competition between ‘be’ and ‘have’ as auxiliaries? Which verbs tend to select ‘be’ rather than ‘have’, and is this selection consistent across different languages?

The distinction between unergative and unaccusative intransitive verbs (Perlmutter 1979, Burzio 1981) has been instrumental in explaining various linguistic properties, among which their ability to be used in impersonal constructions, participial constructions, their compatibility with partitive pronouns (*ne* in Italian and *en* in French), but also their auxiliary selection properties: unergative verbs invariably select ‘have’, whereas unaccusative verbs tend to select ‘be’.

Unergative verbs, characterized by an external subject similar to transitive verbs (Levin and Rappaport Hovav 1995) and typically agents, consistently select the auxiliary ‘have’. For instance, in Italian, *I giornalisti hanno concluso le indagini* ‘the journalists concluded the investigation’ illustrates this pattern where the subject (‘the journalists’) acts upon an object (‘the investigation’). On the other hand, unaccusative verbs feature a subject that undergoes the action rather than initiating it, often described as a patient. For example, in Italian, *nuove prove sono emerse dalle indagini* (‘new evidence emerged from the investigation’), utilizing the verb ‘emerge’ cited as an example of unaccusative verbs in Levin and Rappaport Hovav (1995: 19).

These distinctions are closely tied to telicity, which refers to the inherent completion or endpoint of an action. Telicity is often associated with transitive constructions where the action has a delimited endpoint, because their object is undergoing change with a definite endpoint. For example, ‘John began to write a novel’ (Pustejovsky 1995: 429). Similarly, intransitive verbs of the unaccusative type can exhibit telicity when their subject (corresponding to the role of patient) is conceptualized as undergoing a change or motion with a definite endpoint (e.g., *la pianta è cresciuta rapidamente* ‘the plant grew rapidly’), whereas intransitive verbs of the unergative type are invariably atelic.

In general, “the pattern of auxiliary selection is problematic since with intransitive verbs the selection of *essere* ‘be’ is considered to be a signal of unaccusative status, whereas the selection

of *avere* ‘have’ a signal of unergative status” (Burzio 1986, Perlmutter 1989, C. Rosen 1981, among others” (Levin and Rappaport Hovav 1995: 6). The challenge arises in verbs that blur the distinction between unaccusativity and unergativity, impacting auxiliary selection. While prototypical unaccusative verbs like *arrivare* ‘arrive’ tend to select ‘be’, and prototypical unergative verbs like *lavorare* ‘work’ select ‘have’, others fall in between or resemble unergatives in having an external subject and agent, yet may still select ‘be’ (e.g., *prevalere* ‘prevail’, as also seen in Sorace 2000).

Understanding why certain verbs select different auxiliaries across languages or exhibit mixed profiles within a single language remains a complex challenge. Therefore, some scholars argue for a more nuanced descriptive framework instead of a simple two-way split to accommodate the observed variation (McFadden 2007: 688). A crucial approach is represented by the hierarchy proposed by Sorace (2000). Sorace identifies different lexical classes of verbs and establishes a correlation between the auxiliary chosen (either ‘have’ or ‘be’) and the concepts of agentivity and telicity. This approach moved beyond a simple unaccusative-unergative distinction, suggesting a more gradual differentiation that could vary even within a specific language.

Previous research has uncovered crucial factors affecting the selection of auxiliaries. However, as already noticed by Giancarli (2015), the analysis is mostly based on introspection and acceptability judgments of the possible combination of ‘have’ and ‘be’ with specific verbs, rather than on observation of empirical data. Moreover, the factors identified to account for auxiliary selection appear to be more effective for Italian than for French.

This study aims to determine the importance of parameters influencing auxiliary selection, focusing specifically on verbs allowing the selection of both auxiliaries, the so-called *peripheral verbs* (Sorace 2000), which are the most challenging to analyze. In other words, the study seeks not only to identify the parameters that lead to the selection of ‘have’ and ‘be’ in specific verbs but also to provide a quantitative estimate of actual usage and the relative weight of each factor for different verbs. This will ensure a homogeneous analysis based on the same parameters for all verbs.

To achieve this goal, we have employed both qualitative and quantitative corpus analyses, with the latter utilizing statistical methods. This combined approach allows us to discern the most significant parameters, analyze actual usage patterns, and determine their relative importance.

Our study emphasizes reproducibility and reliability. As detailed in the methodology chapter, sentences were extracted from corpora using precise CQL (Corpus Query Language) queries, which minimizes selection biases. This approach ensures that the data collection process is systematic, reliable, and replicable. A central component of my thesis is the identification of sixteen parameters or factors designed to comprehensively explain auxiliary selection in Italian and French. By doing so, it contributes to a deeper understanding of verb categorization and its implications within linguistic theory.

The thesis is organized into six chapters, each addressing different facets of the research topic.

Chapter 1 provides an overview of previous syntactic and semantic approaches that have addressed this complex topic. It focuses on the most critical perspectives, beginning with syntactic views followed by semantic theories, including the supporting theories and the criticisms they have faced. The unanswered questions in this domain highlight the necessity of reanalyzing this concept.

Chapter 2 outlines the methodology chosen to address these unanswered questions. The approach combines qualitative and statistical corpus analysis of French and Italian. These two languages were selected because they both exhibit verbs with auxiliary alternation, though the auxiliary distribution differs. The combination of these methods, alongside the consideration of the same parameters for all verbs and the internal causation factor, will be pivotal in analyzing auxiliary distribution from a novel perspective.

Chapter 3 presents the results of a corpus analysis of Italian verbs, and explaining the factors that influence auxiliary selection in Italian peripheral verbs. Italian verbs are categorized into full verbs (e.g., *la serata è continuata meravigliosamente* ‘the evening continued wonderfully’) and semi-auxiliary verbs (e.g., *Matilda ha continuato a lavorare duro* ‘Matilda continued to work hard’). For full verbs, semantic parameters such as internal cause and human traits play a significant role in auxiliary selection. In contrast, for semi-auxiliaries, the type of infinitive and internal cause are crucial factors.

Chapter 4 focuses on the corpus analysis for French verbs, showing how some verbs can present both ‘have’ and ‘be’ as auxiliaries when embedded in different constructions. Three categories will be identified (spatial motion verbs, change-of-state verbs, and verbs of appearance), with crucial factors being the construction and the internal cause factor, as well as the human trait for change-of-state verbs.

Chapter 5 focuses on the statistical analysis of Italian verbs and semi-auxiliary verbs, and French verbs. The two methods chosen, CHAID and Random Forest, will be presented, along

with the results of the analysis. The statistical analyses will confirm the hypotheses initially considered in the corpus analysis, highlighting the importance of combining these two approaches. The statistical analysis was conducted with the valuable assistance of Professor Massimo Mucciardi, Associate Professor of Social Statistics at the University of Messina, and Andrea Briglia, Postdoctoral Researcher at Université Grenoble-Alpes.

Chapter 6 provides a comparative analysis of Italian and French verbs, highlighting their differences and similarities in auxiliary selection. This chapter delves into the nuanced variations and shared patterns that influence how auxiliaries are chosen in both languages. It specifically examines the role of semantic factors and syntactic structures in determining which auxiliary—‘have’ or ‘be’—is used in Italian and French. By exploring these factors, Chapter 6 aims to deepen our understanding of verb behavior across different linguistic contexts and contributes to broaden discussions in linguistic theory.

Finally, the conclusion will summarize the results of my analysis and suggest directions for future research.



## CHAPTER 1: PREVIOUS SYNTACTIC AND SEMANTIC APPROACHES

Over the years, the exploration of the relationship between lexical semantics and syntax has gained increasing importance, as highlighted by Levin and Rappaport Hovav (1995: 1). Specifically, investigating how particular semantic meanings are linked to syntactic expressions has become a focal point.

The analysis of auxiliary selection is a critical aspect of this exploration. Indeed, “the relationship between lexical semantics and syntax has received substantial attention in the context of the Unaccusative Hypothesis” (Levin and Rappaport Hovav 1995: 2). This hypothesis, proposed by Perlmutter (1978) and further discussed in section 1.1, posits that intransitive verbs can be divided into two subclasses based on their syntactic configurations, each corresponding to a distinct lexical semantics.

This chapter aims to review the principal theories in the history of linguistics concerning auxiliary selection. It begins with an examination of the most syntactic approaches in sections 1.1 and 1.2, followed by a discussion of the semantic approach in section 1.3. From section 1.4 onwards, the focus shifts to an analysis of the auxiliary split, with particular emphasis on Sorace’s Auxiliary Selection Hierarchy. This analysis includes the presentation of approaches that follow this hierarchy (section 1.5) and those that critique it (section 1.6).

### 1.1 THE UNACCUSATIVE HYPOTHESIS: THE SYNTACTIC APPROACH (PERLMUTTER, BURZIO)

The phenomenon of auxiliary split pertains exclusively to intransitive verbs and hinges on the distinction between two types of intransitive verbs: unaccusative and unergative verbs. The foundational basis for studying unaccusative and unergative verbs is the syntactic Unaccusative Hypothesis (UH), theorized by Perlmutter in 1978 and later extended by Burzio (1986) within the Government-Binding (GB) framework, as noted by Levin and Rappaport Hovav (1995: 2).

Perlmutter’s hypothesis employs the theoretical framework of Relational Grammar, offering a syntactic definition that differentiates between transitive and intransitive verbs, further subdividing the latter into unergative and unaccusative verbs.

In the Government-Binding (GB) approach, as adopted by Levin and Rappaport Hovav, the distinction between unergative and unaccusative verbs is elucidated through their “different underlying syntactic configurations” (Levin and Rappaport Hovav 1995: 3). Specifically, unergative verbs are characterized by the presence of an external argument in the D-structure,

whereas unaccusative verbs are defined by the absence of an external argument, as illustrated by the following schema:

Unergative verb: NP [VP V]  
Unaccusative verb: [VP V NP/CP]

Table 1 - D-Structure syntactic configurations (Levin, Rappaport Hovav 1995: 3)

In unergative verbs, the NP functions as the external argument, while in unaccusative verbs, the NP or CP serves as the internal argument. Alternatively, unaccusatives and unergatives can be differentiated based on syntactic function of the unique argument of the verbal predicate: in unergative verbs, the arguments align with syntactic subjects, and in unaccusative verbs, the arguments align with syntactic objects (Sorace 2000: 879). Verbs are therefore classified as unergatives or unaccusatives based on the nature of their unique argument.

Perlmutter's insight had the significant merit of distinguishing two major classes of verbs, asserting that "the two classes of intransitive verbs are syntactically defined" and using "the difference in syntactic configuration to explain many of the diagnostics that reveal differences in behavior between the classes" (Levin and Rappaport Hovav 1995: 4). Although Perlmutter's approach is syntactic, Levin and Rappaport Hovav (1995: 4) note that "much of the initial research on unaccusativity was directed toward establishing the syntactic aspect of unaccusativity," indicating that semantics is not entirely excluded. In other words, "the Unaccusative Hypothesis was introduced by Perlmutter within the broader context of the Universal Alignment Hypothesis, which suggests that the syntactic expression of arguments can always be determined based on the meaning of the verb" (Levin and Rappaport Hovav 1995: 4). Thus, while the distinction is syntactically represented, it is semantically determined.

## 1.2 UNACCUSATIVE VS. UNERGATIVE VERBS: AT THE INTERFACE BETWEEN SYNTAX AND SEMANTICS (GRIMSHAW, LEVIN & RAPPAPORT)

Most analyses focusing on unaccusative and unergative verbs, as in general for linguistics, locate the phenomenon of auxiliary selection at the interface between syntax and semantics, as noted by Grimshaw (1990) and Levin and Rappaport Hovav (1995).

***Grimshaw's Argument Structure*** Grimshaw (1990) identifies for each verb a lexical-syntactic representation, known as its argument structure. Her work emphasizes that argument structure is not just a list of arguments but a structured representation that reflects the syntactic and semantic relations between the verb and its arguments. This structure is crucial in determining the realization of arguments in syntax.

***Argument Structure, Thematic Roles, and Aspect*** In Grimshaw's view, to account for a complete argument structure, three aspects must be considered: a thematic dimension, an aspectual dimension, and the feature of prominence.

The thematic dimension considers a hierarchy of thematic roles based on the feature of prominence. The thematic hierarchy helps determine which argument takes on the most prominent syntactic roles, such as the subject.

The aspectual dimension anchors the aspectual typology in the hierarchy of argument structure, i.e., "a structured representation which presents prominence relations among arguments" (Grimshaw 1990: 4). It accounts for how arguments are distributed across different events or sub-events.

Prominence, instead, considers the structural organization of the argument array is determined by universal principles based on the semantic properties of the arguments (Grimshaw 1990: 7).

Grimshaw's thematic hierarchy suggests that the agent is typically the most prominent argument both thematically and syntactically since it is realized as the subject. However, the aspectual dimension may be important as she demonstrates with psychological verbs. For example, the *fear*-class verbs present the syntactic subject as the experiencer, while the *frighten*-class verbs present the theme as the subject, which is less prominent than the experiencer.

Grimshaw concludes that it is necessary to consider the aspectual dimension because "the thematic prominence relations of a-structure do not match expectations based on constituent structure" (Grimshaw 1990: 9), as shown with the psychological verbs as an example. The aspectual dimension clarifies that two verbs with similar thematic roles may have different aspectual characteristics. For instance, "arguments of *fear* and *frighten* are thematically the same but they differ fundamentally in that for *frighten* the theme is a cause, and in the *fear* class, it is not" (Grimshaw 1990: 24).

***A-Structure and D-Structure with Unergative and Unaccusative Verbs*** Grimshaw, along with Levin and Rappaport Hovav, posits that unaccusativity is represented syntactically and

sees a-structure as an interface between the semantic and syntactic dimensions of a verb. According to Levin and Rappaport Hovav (1995:1), linking rules are responsible for “the association of arguments bearing certain semantic roles to particular syntactic expressions.”

This allows us to delve into the distinction between unaccusativity and unergativity. In the two-dimensional system, a verb lacks an external argument if no single argument is most prominent in both dimensions. Unaccusative predicates, therefore, must have arguments that fail to reach maximal prominence either in the aspectual dimension or in the thematic dimension (or both) (Grimshaw 1990: 38-39). For unaccusative verbs, the single argument does not fulfill the role of an external argument due to its aspectual dimension. As previously noted, a predicate is seen as a complex event comprising two sub-events: “an argument which participates in the first sub-event in an event structure is more prominent than an argument which participates in the second sub-event” (Grimshaw 1990: 26).

Therefore, “the event structure of an unaccusative corresponds to the second subpart of an accomplishment”, making their unique argument internal, while “the event structure of an unergative corresponds to the first subpart” (Grimshaw 1990: 40). Grimshaw goes further, showing how the semantic explanation behind unaccusatives and unergatives has implications at the syntactic level.

The conclusion is that “the aspectual status of the unaccusative argument makes it internal and places it in d-structure object position, not subject position. The aspectual status of the theme argument of a *frighten* psychological predicate guarantees that it is realized as a d-structure subject. Thus, although both are internal arguments, their d-structure positions are different because their aspectual statuses are different” (Grimshaw 1990: 42).

Consequently, Grimshaw proposes a true hierarchy. Transitives are characterized by two different levels because the semantic information of a verb translates into two different syntactic levels: the subject in the d-structure and the object complement.

Intransitives, based on an aspectual and therefore semantic basis, are characterized by only one level. At the argument structure level, the external argument is present in unergatives, and the internal argument is present in unaccusatives. Syntactically, the external argument is realized as the subject of the d-structure, and the internal argument as the object, as illustrated in Table 1.1.

<b>Grimshaw (1990)</b>			
	Aspectual/Thematic properties	A-structure	D-structure
<b>Transitives</b>	x (agent <sup>1</sup> ) – y (theme)	x (external argument) – y (internal argument)	x (subject) – y (object)
<b>Unergatives</b>	x (agent)	x (external argument)	x (subject)
<b>Unaccusatives</b>	y (theme)	y (internal argument)	y (object)

TABLE 1.1 – Representation of Grimshaw’s theory on the different levels

***Levin and Rappaport Hovav’s Contribution*** In line with Grimshaw (1990), Levin and Rappaport Hovav (1995) also emphasize the importance of the syntax-semantics interface, assuming that semantic information plays a role in selecting the syntactic relations of a verb, or that “certain lexical semantic distinctions are relevant to determining a verb’s argument structure” (Levin and Rappaport Hovav 1995: 136). Levin and Rappaport Hovav (1995: 13-14) argue that “the syntactic classification of verbs is semantically determined. But the fact that the verb classes can be given a semantic characterization does not preclude the attribution of common syntactic properties to all unaccusative verbs.”

In line with Grimshaw (1990), their approach distinguishes a semantic lexical representation and a syntactic lexical representation, the aforementioned argument structure which “encodes the syntactically relevant argument-taking properties of the verb” (Levin and Rappaport Hovav 1995: 20-21), which then has its own representation on the syntactic level. The difference is that there is an enrichment on the lexical level (cf. Van Valin, §1.3, for a similar analysis), in that they take on a lexical representation that has the form of a predicate decomposition (Levin and Rappaport Hovav 1995: 22-23), which can represent the elements common to the verbs’ meanings.

For example, the sentence ‘The window broke’ can be analyzed as follows:

- (1) The window broke → noncausative break : [y **BECOME BROKEN**] or [y **BECOME** predicate]

The sentence in (1) is seen as characterized by the operator **BECOME**, and ‘broken’, which represents the acquired state of being broken (see Levin and Rappaport Hovav 1995: 23 for more).

---

<sup>1</sup> This representation is only of the prototypical roles, but the different thematic roles depend on the verbal class under consideration.

If we take, instead, a causative verb as *dry*, the semantic representation is as follows:

(2) The sun makes the clothes dry → causative dry:

[[ x ACT ] CAUSE [y BECOME dry]]<sup>2</sup> or [[x ACT] CAUSE [y BECOME predicate]

where “BECOME predicate” is the resulting state.

The aspectual dimension is a shared focus in both Grimshaw’s and Levin and Rappaport Hovav’s approaches. However, Grimshaw emphasizes the concept of events and sub-events, exemplified by the verb ‘break’ as a complex event involving two sub-events: ‘activity’ and ‘achievement’, terms also found in Van Valin’s framework.

In contrast, Levin and Rappaport Hovav also address the aspectual dimension, viewing verb meanings as decompositions comprising elements present in both the logical structure and Van Valin’s semantic classes.

### 1.3 UNACCUSATIVE VS. UNERGATIVE VERBS: A SEMANTIC APPROACH (VAN VALIN)

It will be possible to delve into the nuances of split intransitivity only after analyzing the intricate distinction between unergativity and unaccusativity, particularly from a semantic perspective as elucidated by Van Valin (1990) in his theoretical framework of Role and Reference Grammar. Although Van Valin’s framework primarily addresses syntax, it offers a semantic interpretation of the differentiation between unaccusative and unergative verbs, also known as the auxiliary split. This distinction builds upon Vendler’s (1957) semantic typology of verbs, which was further developed by Dowty (1979).

***Vendler’s typology*** Central to Vendler’s (1957: 143) verb classification is the concept of time, as he asserts, “the fact that verbs have tenses indicates that considerations involving the concept of time are relevant to their use.”

Vendler distinguishes four semantic classes: *state*, *activity*, *achievement*, and *accomplishment*, as table 1.2 illustrates.

---

<sup>2</sup> I have used different words to describe the semantic predicate decomposition of Levin and Rappaport Hovav (1998: 107) to represent these concepts in a homogeneously way in this work.

SEMANTIC CLASSES	EXAMPLES
<i>State</i>	e.g. <i>have, love, desire</i>
<i>Activity</i>	e.g. <i>run, swim, walk</i>
<i>Achievement</i>	e.g. <i>reach (the summit), win (the race), cross (the border)</i>
<i>Accomplishment</i>	e.g. <i>paint (a picture), build (a house), write (a novel)</i>

TABLE 1.2 – Vendler’s typology (1957: 150)

Tense can be used in order to differentiate these types of verbs: for example, a continuous tense allows us to differentiate **Activity verbs** from Achievement verbs. *Run* is regarded an activity verb because it represents an *ongoing process*, so much that “if someone has been running for half an hour, then it must be true that he has been running for every period within that half-hour [...]. Running and its kind go on in time in a homogenous way; any part of the process is of the same nature as the whole” (Vendler 1957: 146).

**Verbs of accomplishment**, on the other hand, “also go on in time, but they proceed toward a terminus which is logically necessary to their being what they are”, as for *write a novel*. The presence of the object complement in the examples is not hazardous but makes us understand how some verbs can belong to more than one category. Both the verbs *write* and *run*, in fact, are verbs of activity that can become verbs of accomplishment in the presence of a *terminus* (as in *run a mile* or *write a novel*).

**Achievement verbs** describe an action that “occurs in a single moment” (Vendler 1957: 147), so much that, as the reach the summit example shows, the summit can be reached at a given time, but “even if one says that it took him three hours to reach the summit, one does not mean that the reaching of the summit went on during those hours. Obviously, it took three hours of climbing to reach the top” (Vendler 1957: 147).

And finally, continuous tenses can also help distinguish **State verbs** from the other semantic classes, like activity. More precisely, “the lack of continuous tenses (e.g., “I am knowing, loving, and so forth”) is enough to distinguish them from activities and accomplishments, and the form of determination (“How long...?” “For such-and-such a period”) should be sufficient to keep them from being confused with achievements” (Vendler 1957: 148).

**Vendler’s typology revisited by Dowty (1979)** Dowty (1979) revisits Vendler’s verb classes from a logical perspective, offering predicate decompositions that involve state, achievement, activity, and accomplishment.

His analysis introduces logical operators like BECOME and CAUSE, defining how these verbs are structured semantically, as depicted in table 1.4.

VERB CLASS	PREDICATE DECOMPOSITION <sup>3</sup>	EXAMPLE	PREDICATE DECOMPOSITION
State	predicate' (x) or (x,y)	<i>The book is heavy.</i>	be' (book, [heavy'])
Achievement	<b>BECOME</b> predicate' (x) or (x,y)	<i>Susan arrived at the house.</i>	<b>BECOME</b> be-at' (house, Susan)
Activity (Agentive)	(± <b>DO</b> (x)) [predicate' (x) or (x,y)]	<i>Susan ran.</i>	run' (Susan)
Accomplishment	<sup>4</sup> [activity] <b>CAUSE</b> [achievement]	<i>Linda threw the hat on the table.</i>	[throw' (Linda, hat)] <b>CAUSE</b> [ <b>BECOME</b> be-on' (table,hat)]

TABLE 1.4 – Dowty (1979)'s revision of Vendler's typology presented by Van Valin (1990: 224)

Van Valin (1990) considers Dowty's revised version of Vendler's verb classes. Each predicate decomposition, hence, each verb structure, comprises a predicate and one or two variables. Many of these structures involve logical operators such as DO, BECOME, or CAUSE, either individually or in combination.

The operator DO distinguishes states from activities. Dowty (1979) explains that activities derived from stative predicates are states that a person can voluntarily enter and maintain, such as 'seeking answers' or 'being polite', whereas 'being in Chicago', 'being blond', and 'knowing the answers' are not activities in this sense. Dowty argues that DO should not be reduced to mere intentionality but rather involves an implicature based on the animacy of the actor and lexical properties of the verb (Van Valin 1990: 224).

In the context of inchoative verbs, which primarily characterize achievement verbs, BECOME serves as the appropriate logical operator (Dowty 1979: 77). Furthermore, Dowty notes that accomplishment verbs typically involve two phrases of different types. Generally, these verbs consist of an activity and an achievement connected by the CAUSE operator, resulting in sentences of the following form:

<sup>3</sup> I replaced the term *logical structure* by *predicate*, in order to arrive at a homogeneous terminology in relation to the other approaches presented here.

<sup>4</sup> " $\varphi$  CAUSE  $\psi$  where  $\varphi$  is normally an activity predicate and  $\psi$  an achievement predicate", in the original text, is replaced here by "[activity] CAUSE [achievement]" for the purpose of clarification.

(3) [[Linda throws the hat] CAUSE [BECOME a hat is on the table]]

These representations are significant as they form the foundation of the RRG theory of semantic roles, which posits two tiers of semantic roles (Van Valin 1990: 225). The first tier consists of universal thematic relations, while the second tier comprises two macro-roles: the Actor and the Undergoer. Thematic relations are established based on predicate decomposition.

In transitive verbs, the Actor and Undergoer are the two primary arguments, whereas in intransitive verbs, the sole argument can function as either an Actor or an Undergoer. These macro-roles encompass various specific thematic relations; for instance, Actor can denote an agent, effector, experiencer, and more, while Undergoer can represent a patient, location, theme, or even an experiencer. The Actor-Undergoer Hierarchy delineates the hierarchical relationship between these classical semantic roles concerning the macro-roles Actor and Undergoer (Figure 1.5).

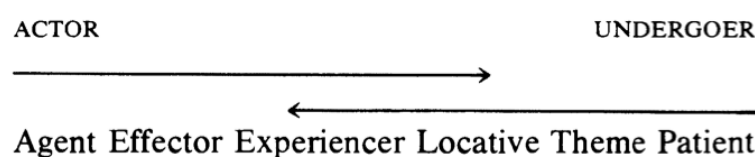


FIGURE 1.5 – Actor-Undergoer Hierarchy (Van Valin 1990: 226)

**Van Valin's (1990) hypothesis** Previous approaches have elucidated the characteristics of unaccusativity and unergativity but have not explicitly explained why 'have' is associated with unergativity and 'be' with unaccusativity. This issue becomes more intricate when considering the selection of either auxiliary with the same verb.

The treatment of auxiliaries is not central in Grimshaw (1990), as these concepts are explored within a broader framework devoted to argument structure. Levin and Rappaport Hovav (1995) delve deeper into this topic, although it remains part of the overarching debate on unaccusativity versus unergativity, primarily focused on English.

In contrast, Van Valin (1990) offers a purely semantic rationale for auxiliary selection in his RRG framework, using Italian as an illustrative example where intransitives present a complex scenario. Verbs that employ the auxiliary *avere* 'have' in perfect tenses are termed 'A-verbs', derived from the auxiliary *avere*, while those using *essere* 'be' are termed 'E-verbs', derived from the auxiliary *essere*.

According to Van Valin, the distinction hinges on *Aktionsart*, namely the inherent aspectual properties of the verbs. Van Valin (1990: 232) contrasts activity verbs with state, achievement,

and accomplishment verbs. Activity verbs are categorized as A-verbs, while state, achievement, and accomplishment verbs are classified as E-verbs.

For instance, a verb like *parlare* ‘talk’ is categorized as an activity verb because its logical form involves the action of talking and selects ‘have’ in the perfect tense. Conversely, a verb like *arrivare* ‘arrive’ is deemed an achievement verb because it brings the subject to a specific state indicated by the predicate, thereby selecting ‘be’ in the perfect tense. Van Valin highlights that E-verbs typically feature a state predicate in their predicate decomposition.

The following table demonstrates the distribution of different aspectual classes concerning the distinction between E-verbs and A-verbs:

	VERB CLASS	LEXICAL DECOMPOSITION	EXAMPLE
<b>A-VERBS</b>	<i>Activity</i>	(DO (x,)) [ <b>predicate</b> ’ (x) or (x,y)]	e.g. <i>ballare</i> ‘dance’ → <b>dance</b> ’ (x)
<b>E-VERBS</b>	<i>State</i>	predicate’ (x) or (x,y)	e.g. <i>stare</i> ‘stay’ → <b>be-at</b> ’ (x)
<b>E-VERBS</b>	<i>Achievement</i>	BECOME predicate’ (x) or (x,y)	e.g. <i>arrivare</i> ‘arrive’ → <b>BECOME be-at</b> ’ (x)
<b>E-VERBS</b>	<i>Accomplishment</i>	[activity] CAUSE [achievement]	e.g. <i>andare</i> ‘go’ → [do’ (x)] CAUSE [ <b>BECOME be-at</b> ’ (y,x)]

TABLE 1.6 – Aspectual verb classes and auxiliary selection according to Van Valin (1990: 233)

Complications arise for verbs that can appropriately take both ‘have’ and ‘be’. An illustrative example is provided by the verb *saltare* ‘jump’:

- (4) *Ida ha saltato (di proposito) sul letto.*  
 Ida have. PRS.3SG jump.PTCP of.purpose on.DEF.ART bed

‘Ida has jumped [up and down] on the bed (on purpose).’

- (5) *Ida è saltato (di proposito) dalla finestra.*  
 Ida be. PRS.3SG jump.PTCP.F.SG of.purpose from.DEF.ART window

‘Ida has jumped out of the window (on purpose).’

(Van Valin 1990: 235)

According to Van Valin, the choice between the auxiliary verbs ‘have’ and ‘be’ with A/E motion verbs is directly related to whether the verb expresses activity or accomplishment

semantics (Van Valin 1990: 237). Therefore, the alternation of auxiliaries for the same verb reflects an aspectual difference: for instance, *Ida ha saltato sul letto* (4) represents an activity, while *Ida è saltata dalla finestra* (5) represents an accomplishment.

Sometimes, especially with intransitive motion verbs, the distinction hinges on whether the motion is bounded by a specific starting or ending point (Van Valin 1990: 236). For example, *sul letto* ‘on the bed’ suggests ongoing activity, whereas *dalla finestra* ‘from the window’ indicates a completed action with a clear starting and endpoint.

It is pertinent to question whether there exists a semantic difference between *Ida ha saltato sul letto* (4) and *Ida è saltata sul letto* (5). Syntactically, these sentences are identical except for the auxiliary used. If we posit a semantic distinction between them, we must consider what triggers the selection of ‘have’ or ‘be’.

Regarding the verb *correre* ‘run’, it is often noted that the presence of an endpoint can influence the choice of auxiliary:

- (6) Ugo ha corso meglio ieri.  
 Ugo have. PRS.3SG run.PTCP better yesterday

‘Ugo ran better yesterday.’

- (7) Ugo è corso a casa.  
 Ugo be. PRS.3SG run.PTCP to home

‘Ugo run home.’

(Levin and Rappaport Hovav 1995: 186)

According to Levin and Rappaport Hovav’s theory, the choice between the two auxiliaries ‘have’ and ‘be’ depends on a different semantic representation, particularly influenced by the presence of a directional phrase. This distinction applies to verbs that denote manner of motion and exhibit directed motion use in Italian (Levin and Rappaport Hovav 1995: 185).

For instance, the verb *correre* ‘run’ is categorized as unergative when expressing manner of motion and shifts to unaccusative when its lexical semantic representation includes a directional phrase.

Consider sentence (8): it presents a challenge in understanding why the auxiliary *essere* ‘be’ is used, as there is neither a specified endpoint nor an element that indicates a resultative action or an accomplishment sentence.

(8) Luisa è corsa via di fretta.  
 Luisa be. PRS.3SG run.PTCP.F.SG away in.a.hurry

‘Luisa ran away in a hurry.’

Certainly, analyzing the interplay between semantics and syntactic contexts is crucial to comprehending auxiliary selection in languages like Italian. Rather than viewing the auxiliary purely as a syntactic element, it carries significant semantic implications such as temporality and aspectuality. For instance, applying Grimshaw’s approach to sentences like *Luisa ha corso nel parco* (‘Luisa has run in the park’) and *Luisa è corsa nel parco* (‘Luisa has run in the park’), we observe that the choice between the auxiliaries *ha* ‘have’ and *è* ‘be’ reflects different aspectual semantics. *Luisa è corsa nel parco* conveys a resultative semantics, where the choice of *è* ‘be’ aligns with this specific aspectual interpretation. This suggests that the semantics of a verb isn’t predetermined but rather influenced by the syntactic context in which it appears.

Levin and Rappaport Hovav (1995: 5-6) argue against the notion that semantics alone can explain phenomena like unaccusativity, particularly referencing Rosen’s (1984) critique of the ‘Little Alignment Hypothesis’ (see Levin 1995: 288-289). This hypothesis posits a fixed mapping between semantic roles and grammatical relations, suggesting that each verb adheres to a specific pattern in aligning its semantic roles with syntactic structures. However, Rosen disputes this by considering three aspects.

First, some verbs exhibit variability in auxiliary selection, challenging the idea of a fixed mapping. Indeed, the same verb might select different auxiliaries depending on subtle shifts in meaning or context.

Second, Rosen’s critique indicates that verbs cannot be neatly classified based solely on their semantic properties because they can demonstrate different behaviors in different contexts: “verbs with similar meanings in and across languages may be classified differently with respect to unaccusativity” (Levin 1995: 6).

Finally, there is no semantic property common to all unaccusative verbs across all languages as a diagnostic. This suggests that verbs may possess multiple semantic representations that share a core meaning but differ in aspects relevant to classifying them as unaccusative or unergative.

This perspective underscores the complexity of auxiliary selection and the need to consider both syntactic and semantic factors. It suggests that a purely semantic approach is insufficient and that a more integrated analysis, considering how syntax and semantics interact, is necessary to fully understand phenomena like unaccusativity. By acknowledging the variability in verb behavior, Levin and Rappaport Hovav highlight the importance of a nuanced approach that

takes into account the dynamic interplay between a verb’s semantic roles and its syntactic expression.

*Van Valin, Grimshaw and Levin and Rappaport Hovav: a comparison* In Role and Reference Grammar (Van Valin 1990), unlike previous approaches such as Relational Grammar (RelG) and Government and Binding theory (GB), the distinction between unaccusative and unergative verbs is primarily semantic. This distinction is rooted in both thematic relations and aspectual considerations.

When comparing this semantic approach with the syntactic perspectives proposed by Grimshaw (1990) and Levin and Rappaport Hovav (1995), there are notable differences but also some similarities, as illustrated in Table 1.7.

**UNACCUSATIVITY VS. UNERGATIVITY**

	<b>VAN VALIN (1990)</b>	<b>GRIMSHAW (1990)</b>	<b>LEVIN AND RAPPAPORT HOVAV (1995)</b>
<b>LEXICAL SEMANTIC REPRESENTATION</b>	Aspectual Dimension (Predicate decomposition)	Aspectual Dimension (Subevents)	Aspectual Dimension (Predicate Decomposition) <sup>5</sup>
	Thematic Dimension (thematic relationship + Actor/Undergoer hierarchy)	Thematic Dimension (Thematic Hierarchy)	
<b>LEXICAL SYNTACTIC REPRESENTATION</b>	/	A-Structure (interface semantic information - syntactic realization)	A-Structure (linking rules responsible of the A-Structure)
	Unacc. and Unerg. semantically explained	Unacc. and Unerg. semantically determined, syntactically encoded	Unacc. and Unerg- semantically determined, syntactically encoded

TABLE 1.7 – The approaches of Van Valin, Grimshaw and Levin and Rappaport Hovav: a comparison

<sup>5</sup> The absence of the thematic dimension in this table does not mean that thematic roles are not considered at all. Here I want to emphasize the role that this dimension plays with respect to unergativity and unaccusativity. In Grimshaw the importance of the dual dimension has some relevance to the phenomenon described.

Regarding lexical semantic representation, the aspectual dimension is crucial for all three approaches, viewed as lexical decomposition in both Van Valin (1990) and Levin and Rappaport Hovav (1995).

In contrast, the thematic dimension is only conceptualized by Grimshaw (1990) and Van Valin (1990), albeit with some distinctions: Van Valin (1990) focuses on the actor/undergoer hierarchy and thematic relationships, whereas Grimshaw (1990) emphasizes the thematic hierarchy. Van Valin's approach (1990) is inherently semantic, whereas Grimshaw (1990) and Levin and Rappaport Hovav (1995) consider the presence of an Argument Structure from both lexical and syntactic viewpoints.

Thus, while Van Valin (1990) semantically explains unaccusativity and unergativity, Grimshaw (1990) and Levin and Rappaport Hovav (1995) recognize the importance of semantics in determining unaccusativity and unergativity but also emphasize their syntactic encoding.

#### 1.4 AUXILIARY SELECTION HYPOTHESIS (SORACE)

For Romance languages, the subject of auxiliary selection has been addressed extensively by Sorace (2000), whose work plays a pivotal role. Sorace (2000) states:

“Auxiliary selection is sensitive to a hierarchy of aspectual/thematic verb types: some verbs require a given auxiliary categorically, whereas others allow both auxiliaries to a greater or lesser extent depending on their position of the hierarchy. It is argued that this gradience has potentially important implications for the UNACCUSATIVE HYPOTHESIS, and more generally for theories of the lexicon-syntax interface.”

The gradience of this alternation is central to her approach, which, as we will show below, is fundamentally semantic. This gradience led Sorace to propose a hierarchy that has been adopted by many other scholars (see, among others, Bentley and Eythórsson 2003, McFadden 2007). The idea is that verbs can be semantically divided into different classes, and that for each class, it is possible to state how much variation is expected.

This proposal is called the Auxiliary Selection Hierarchy (ASH), “a more nuanced descriptive framework than a simple two-way split” (McFadden 2007), as depicted in Table 1.8.

Change of Location	<b>BE</b>
Change of State	
Continuation of a pre-existing state	
Existence of a State	
Uncontrolled Process	
Controlled Process (Motional)	
Controlled Process (Non-motional)	<b>HAVE</b>

TABLE 1.8 – The Auxiliary Selection Hierarchy (Sorace 2000: 863)

To understand this table, it is important to know that two factors have been considered essential for the hierarchy: telicity and agentivity. Telicity correlates with ‘be’ and agentivity with ‘have’, “whose interaction affects the syntax of split intransitivity and creates gradient satisfaction of morphosyntactic diagnostics of split intransitivity” (Sorace 2015: 24). More fundamentally, at the basis of this hierarchy, particularly related to intransitive verbs, there is the assumption that each predicate may be “associated with an event-structure template that can be conceptualized in terms of two distinct aspectual subevents: a process, on the one hand, and a transition or state, on the other” (Sorace 2000: 862).

Sorace aligns with the work of Grimshaw (1990), insofar as an accomplishment verb can be seen as the result of two events: an activity (according to Sorace rather a “process”, *i.e.* dynamic, without presupposing an agent) and a state. The second part of the hierarchy correlates more with verbs that would belong to Grimshaw’s first sub-event. If we examine them in more detail, Sorace claims that three more criteria divide verbs denoting processes: the nature of the causation of the process, affectedness, and density. In contrast, the first part of the hierarchy contains verbs that we might define as belonging to Grimshaw’s second sub-event.

For Sorace (2000: 863), their semantic lexical representations correspond not only of states but also to change of states. These verbs can be characterized by two dimensions: telicity, already mentioned, and the degree of dynamicity. For instance, the higher the degree of dynamicity and telicity in a verb, the more likely the auxiliary is ‘be’. Verbs of change of location, like the Italian *venire* ‘to come’, always select ‘be’. Verbs lower on the hierarchy are expected to exhibit more variation.

In contrast with Van Valin’s approach, Sorace’s hierarchy is built on the concept of gradience: some verbs more likely correlate with ‘be’ and others with ‘have’, but the cut-off point differs according to languages, which may explain why there could be different auxiliary selection cross-linguistically. Indeed, the concept of gradience “is a reflection of distinction

among intransitive verbs in terms of their aspectual and thematic structure, which define a hierarchy of verbs that are likely to prefer one or the other auxiliary” (Sorace 2000: 886) or may allow both auxiliaries to different degrees. According to this hierarchy, some intransitive verbs categorically select ‘be’ or ‘have’; others select ‘be’ in certain languages and ‘have’ in others, and finally, some verbs show a high degree of variation, reflected in different degrees of acceptability or unacceptability judgments from the speakers.

Examples (9) and (10) show that change-of-location verbs in French and Italian both select the auxiliary ‘be’, regardless of the feature of agentivity:

- (9) Maria è arrivata in ritardo. (Italian)  
 Maria be. PRS.3SG arrive.PTCP.F.SG late

‘Maria arrived late.’

- (10) Marie est arrivée en retard. (French)  
 Maria be. PRS.3SG arrive.PTCP.F.SG late

‘Maria arrived late.’

Sorace (2000: 863)

Indeed, the verbs *arrivare* and *arriver* ‘come’ are at the highest point of the hierarchy and strongly select ‘be’, without any hesitation. In the case of change-of-state, ‘be’ is used for Italian (11) and ‘have’ (or, less likely, ‘be’) for French (12):

- (11) Maria è cambiata. (Italian)  
 Maria be. PRS.3SG change.PTCP.F.SG

‘Maria has changed.’

- (12) Maria a (est) changée. (French)  
 Maria have(be). PRS.3SG change.PTCP.F.SG

‘Maria has changed.’

(11) and (12) demonstrate that as one moves further down from the change-of-location category at the top of the hierarchy in Table 1.8, the occurrence of ‘have’ becomes more frequent, not only within the same language but potentially cross-linguistically as well.

The situation becomes even more complex when variations in auxiliary selection occur within the same verb in a given language, like in (12) and, as previously noted with the Italian verb *correre* ‘run’, which exhibits sensitivity to telicity. For instance, data from Sorace (2000:

876) indicates that a telic aspectual interpretation causes the verb *correre* to undergo an auxiliary shift in Italian, whereas this phenomenon does not occur in French.

***The verb’s classes of Sorace’s ASH*** We will now offer an overview of the different semantic categories established by Sorace, and listed in Table 1.8: change of location, change of state, continuation of a pre-existing state, existence of a state, uncontrolled process, controlled process (motional), and controlled process (non-motional).

***Change-of-location verbs*** (e.g., *venire* ‘come’, *cadere* ‘fall’) involve a concrete displacement from one point in space to another (Sorace 2000: 863). These verbs show minimal variability in Italian and French: “one important feature of change-of-location verbs is that they select ‘be’ regardless of other aspectual features contributed by the sentence in which they appear. The detelicization of the predicate, for example, does not impact auxiliary selection [...]. Similarly, agentivity does not affect auxiliary choice with these verbs” (Sorace 2000: 864). Therefore, inherently telic verbs choose ‘be’ independently of other factors affecting the sentence.

***Change-of-state verbs*** (e.g., *salire* ‘rise’, *diventare* ‘become’) express a change in a particular direction without specifying a telic endpoint (Sorace 2000: 864); there are exceptions such as ‘die’ and ‘be born’. There appear to be two types of change-of-state verbs: those that tend to trigger an interpretation in which the *telos* is reached (e.g., ‘decay’), and those that favor an interpretation in which there is a movement toward a *telos* without however reaching it (e.g., ‘rise’) (Sorace 2000: 865). This class of verbs is generally considered less telic than change-of-location verbs. ‘Be’ is preferred in this class of verbs, as seen with the verb *salire* ‘rise’, but variation between ‘have’ and ‘be’ is not excluded, as with the verb *fiorire* ‘blossom’.

***Continuation-of-condition verbs*** (e.g., *restare* ‘stay’, *durare* ‘last’) denote the continuation of a preexisting condition (Sorace 2000: 867). In Italian, ‘be’ seems more appropriate, but in some cases, ‘have’ is still acceptable. Sorace (2000: 867-868) provides examples where the verb *durare* ‘last’ can take both ‘have’ and ‘be’. Sentences with a [+HUMAN] subject show a more acceptable use of ‘have’:

- (13) La guerra è/?ha durato a lungo.  
 DEF.ART war be/?have. PRS.3SG last.PTCP for.long

‘The war lasted a long time.’

- (14) Il presidente è/ha durato in carica due anni.  
 DEF.ART president be/have. PRS.3SG last.PTCP in.post two years

‘The president lasted in post for two years.’

(Sorace 2000: 868)

According to Sorace, “the contrasts show that verbs denoting continuation of state, unlike verbs of change of location or condition, are sensitive to the agentivity of the subject: agentive subjects make the use of auxiliary *avere* more acceptable” (Sorace 2000: 868).

As to French, it should be noted that agentivity does not seem to influence auxiliary selection: *rester* is conjugated with ‘be’, whereas *durer* takes the auxiliary ‘have’, regardless of the nature of the subject.

**Existence of state verbs** (*essere* ‘be’, *esistere* ‘exist’) is a category subject to variation. These verbs refer to concrete states and abstract or psychological states, like *sembrare* ‘seem’. Generally, these verbs are more likely to select ‘be’ in Italian, but variation is more likely to occur, and in some cases, ‘have’ is acceptable. In French, “verbs of existence consistently select auxiliary avoir” (Sorace 2000: 869).

**Uncontrolled process verbs** (e.g., *tremare* ‘tremble’, *tentennare* ‘waver’) can refer to involuntary bodily processes or external processes involving intrinsic non-agentive subjects. More precisely, they consist of “uncontrolled action, involuntary bodily function, and emission (of substance/light/sound/smell), whose common denominator is the lack of volitionality” (Sorace 2000: 877). Sorace states that “uncontrolled process verbs show more variable behavior” (Sorace 2000: 869). For example, verbs like *tentennare* in Italian can select ‘have’, but ‘be’ is also possible, as in (15):

- (15) La fede religiosa ha/è tentennato/a [...].  
 DEF.ART faith religious have/be. PRS.3SG waver.PTCP/PTCP.F.SG

‘The religious faith wavered [...].’

- (16) Paolo ha/\*è tentennato a lungo [...].  
 DEF.ART have/\*be. PRS.3SG waver.PTCP for.long

‘Paolo wavered for long [...].’

(Sorace 2000: 877)

Sorace (2000: 877) states that “verbs of nonvolitional process show a marked sensitivity to the animacy of the subject: although many can be used with both human and inanimate subjects, they prefer ‘have’ with agentive subjects”, as shown in (16).

The influence of agentivity is evident, as Sorace (2000: 877) demonstrates by noting that verbs denoting involuntary bodily functions (e.g., ‘cough’) can be construed as implying volition more easily, despite typically having human non-agentive subjects. Sorace posits that non-volitional verbs often have animate subjects, while involuntary bodily function verbs involve non-agentive subjects.

This apparent contradiction lies in the expectation that non-volitional verbs should favor ‘be’ due to their non-agentive nature. However, the resolution is found in recognizing the significant role of animacy in auxiliary selection. Animacy influences auxiliary selection, allowing even non-agentive verbs with animate subjects to favor ‘have’.

Another hypothesis that we could consider, in line with Levin and Rappaport Hovav’s approach, is that Sorace, when using the label ‘agentive subject’, means rather ‘causative subject’. For example, Levin and Rappaport Hovav (1995: 137) say that verbs like ‘cough’ “can still be considered internally caused” and thus have an external subject like unergatives.

***Controlled (motional) process verbs*** (e.g., *correre* ‘run’) show more variable behavior. The verb *correre* ‘run’, for instance, as shown above, can take both ‘have’ and ‘be’ in Italian (Sorace 2000: 872). Agentivity appears to have a strong influence on auxiliary selection for these verbs. The more agentive the verb, the more likely it will take ‘have’.

***Controlled (non-motional) process verbs*** (e.g., *lavorare* ‘work’) strongly prefer ‘have’ in both French and Italian, as they denote activities with a high degree of agentivity and typically lack telicity. Among these verbs, the author includes *giocare* ‘play’ and *parlare* ‘talk’. For some of these verbs, like *funzionare* ‘function’, “some native speakers of Italian find the auxiliary *essere* not completely unacceptable when the subject is nonagentive” (Sorace 2000: 874).

Controlled non-motional process verbs are a subclass of process verbs, along with controlled motional process verbs and controlled affecting process verbs. The subclass of controlled affecting processes, on the other hand, includes verbs “denoting volitional actions that imply a permanent change of state for the subject argument, such as *abdicare* ‘abdicate’, *aderire* ‘join in’, *prevalere* ‘prevail’ [...]” (Sorace 2000: 874). The subject of these verbs is strongly agentive and affected, and when it is not human (and, as will be shown by my empirical data, perhaps even when it is human), these verbs could take the auxiliary ‘be’.

***Concluding remarks on ASH*** Considering the hierarchy as a whole, it becomes apparent that two distinct trends converge towards a midpoint. At the top of the hierarchy, telicity is strong and gradually diminishes as one moves down the ordered list of semantic categories. Conversely, at the bottom of the hierarchy, agentivity is strong and diminishes as one moves up the ordered list of verb semantic categories. Telic verbs typically favor ‘be’ as their auxiliary, while agentive verbs are more inclined to select ‘have’.

The question remains whether this hierarchy can effectively explain auxiliary shifts for the so-called ‘midpoint’ in a simpler way.

Furthermore, a certain imbalance between the parameters of agentivity and telicity merits consideration. Verbs situated in the upper segment of the hierarchy demonstrate greater sensitivity to telicity than to agentivity. Agentivity, therefore, might play a role primarily within the lower segment of the hierarchy, although its exact influence remains unclear. However, telicity appears relevant even in the lower segment, where intrinsically atelic verbs behave akin to unaccusatives, insofar as they select ‘be’ in telic contexts. The ASH framework does not clarify the relative weight of various factors, i.e. not only the telicity-agentivity dichotomy mentioned above but also the interplay between animacy and agentivity.

Other issues are the difficulty in delineating boundaries between verb classes and determining which verbs unequivocally belong to a specific class, as well as the unclear motivation for the position of each verb class in the ordered list. For instance, some verbs, such as *prevalere* ‘prevail’, typically prefer ‘have’ but can accept ‘be’ when the subject is non-human (as evidenced in this study, where ‘be’ occurs even with a human subject). Such verbs are positioned close to the unergative end of the hierarchy, without adequately considering that their occasional use of ‘be’ suggests they are not distant from the midpoint of the hierarchy.

Hence, the question arises as to what extent such a hierarchical framework can effectively address a topic as intricate as auxiliary selection, which necessitates consideration of numerous factors – some of which can be language-specific – that challenge the prospect of presenting a uniform and comprehensive view.

#### 1.5 A BINARY FEATURE APPROACH OF THE ASH (BENTLEY & EYTHÓRSSON.)

Sorace’s Auxiliary Selection Hierarchy (ASH) has significantly influenced research on auxiliary selection and related topics, including unaccusativity. Bentley and Eythórsson (2004) build on Sorace’s framework, in arguing that the variation in auxiliary selection is not purely

syntactically motivated but reflects semantic distinctions that underpin the concept of unaccusativity.

Bentley and Eythórsson, whose research is based on a cross-linguistic analysis of Romance and Germanic languages, propose that unaccusativity is semantically determined, with its defining properties organized hierarchically from core to periphery along a typicality scale, and put forward an implicational hierarchy of unaccusative verbs.

According to Bentley and Eythórsson (2004: 456), their study based on Sorace’s insight that unaccusativity has to be thought in gradient rather than binary terms. They rely on the verb classes defined in Sorace’s hierarchy, and argue that these verb classes exhibit predictable patterns in auxiliary selection across languages.

Bentley and Eythórsson argue that the absence of the auxiliary ‘be’ in unaccusative constructions does not negate their syntactic status but rather indicates a different morpho-syntactic realization of tense/aspect values. Specifically, whether a verb selects ‘be’ or ‘have’ depends on its semantic features, represented by {F}, including the features of dynamicity, telicity, stativity.

However, these features may vary cross-linguistically. {Fn} represents a subset of {F} relevant to ‘be’ selection within a specific language context (Bentley & Eythórsson 2004: 460), as represented by the following rule (rule 31 in Bentley & Eythorsson 2004: 463):

If V is marked [+Fn] > ‘be’ + past participle  
Elsewhere > ‘have’ + past participle

For instance, according to this rule, in Italian, verbs selecting ‘have’ do not exhibit the properties {Fn}, viz. non-telic, agentive, and non-stative. An example given is the verb *marciare* ‘march’ (Bentley & Eythórsson 2004: 461).

The {F} properties, namely dynamicity, telicity and stativity, are ranked in Table 1.9.

1	↓	+ Dynamicity + Telicity (+ D + T) >>
2		+ Dynamicity – Telicity (+D – T) >>
3		+ Stativity – Dynamicity (+St – D) >>
4		+ Stativity (+St)

TABLE 1.9 – Table representing Ranking Properties 1,2 and 3 (Bentley & Eythórsson 2004: 461-463)

According to Bentley and Eythórsson (2003: 461-463), verbs marked as [+D] (dynamic) are more likely to select the auxiliary ‘be’ than verbs marked as [+St] (stative). For example, in both Italian and French, verbs belonging to the change of location category (e.g., Italian *andare*, French *aller* ‘go’) select ‘be’. In contrast, verbs denoting existence of a state (e.g., Italian *esistere*, French *exister* ‘exist’) select ‘be’ in Italian but ‘have’ in French (Bentley & Eythórsson 2003: 461).

Dynamic predicates can further be categorized as telic or non-telic. Telicity, an aspectual property, influences the choice of auxiliaries particularly in verbs denoting change of state, as illustrated in the following sentences:

- (17) Ha saltato per evitare il fosso,  
 have. PRS.3SG jump.PTCP to avoid.INF DEF.ART ditch

‘S/he has jumped to avoid the ditch.’

- (18) È saltato un bottone.  
 be. PRS.3SG jump.PTCP INDF.ART button

‘A button has come off.’

(Bentley, Eythórsson 2003: 462)

Bentley and Eythórsson (2003) propose that the telic change of state in the second sentence influences its semantic interpretation, contrasting with the ongoing activity denoted in the first sentence. They use the semantic features of dynamicity vs stativity and telicity for the characterization of Sorace’s verb classes, as shown in Table 1.10:

SORACE’S VERB CLASSES	SEMANTIC PROPERTIES THEORIZED BY BENTLEY AND EYTHÓRSSON
(i) Change of location	+D, +T
(ii) Change of State	+D, (+/-T)
(iii) Continuation of a pre-existing state	-D, +St
(iv) Existence of a state	+St

TABLE 1.10 – Correspondence of Sorace’s verb classes and semantic properties (Bentley & Eythórsson 2004: 463)

Auxiliary selection varies across languages; for instance, Italian exhibits a broader use of ‘be’ compared to French. Therefore, combining these semantic properties not only explains

auxiliary selection (Bentley and Eythórsson 2003: 469) but also helps describe cross-linguistic variations.

For example, “in Italian ‘be’ is selected by verbs of all the classes in (i). In German ‘be’ is selected by verbs of class (i) (e.g., *gehen* ‘go’) and class (ii) (e.g., *sterben* ‘die’, [+D +T], and *wachsen* ‘grow’, [+D -T]), but not by verbs of classes (iii) and (iv) (e.g., *dauern* ‘last’ and *existieren* ‘exist’). In French ‘be’ is selected by all the members of class (i) (e.g., *aller* ‘go’) and the [+D +T] verbs of class (ii) (e.g., *mourir* ‘die’). ‘Have’ is selected by verbs of the other semantic classes” (Bentley and Eythórsson 2003: 464).

While Bentley and Eythórsson (2003) delve into Sorace’s Auxiliary Selection Hierarchy (ASH) to explore the semantic properties of certain verb classes, their primary focus is on a selective group of verbs, particularly manner of motion verbs. The study frequently utilizes examples such as *correre* ‘run’ and *saltare* ‘jump’, which limits the generalizability of their findings. Consequently, questions remain regarding verbs that alternate between selecting both auxiliaries, leaving some areas unexplored.

In conclusion, while the properties proposed align with Sorace’s semantic categories, they may not universally represent accepted categories, as discussed further in section 1.6.

## 1.6 LIMITS OF THE ASH APPROACH ACCORDING TO GIANCARLI (2015)

The ASH has been widely acclaimed as “today’s most elaborate gradient representation dedicated to the auxiliary selection of BE or HAVE within intransitive verbs” (Giancarli 2015: 79).

However, divergent views persist. Giancarli (2015: 79) critiques the essence of Sorace’s approach, particularly emphasizing the limits of graduality and proposing instead a non-gradient and non-scalar perspective. Moreover, he argues for an integration of reflexive verbs into the analysis, which were not considered by previous studies.

Giancarli distinguishes between graduality and scalarity in qualitative and quantitative terms. Scalarity refers to a purely quantitative scale where elements differ only in quantity, whereas graduality encompasses both qualitative and quantitative differences. It relates closely to the notion of a prototype, where some elements are core and most representative of the prototype, while others are peripheral. The concept of graduality is central to the ASH, described as bi-dimensional rather than mono-dimensional, encompassing two classes of elements and an intermediate zone.

Giancarli (2015: 82-86) identifies five key limitations of the ASH framework and its concept of graduality. These issues listed in Table 1.11 summarize Giancarli’s critique of ASH and constitute a plea for a revised framework that can address these issues more effectively.

<b>LIMITS OF ASH</b>	
(1) Agentivity and telicity at the opposites: the ASH scale reduces unergativity to agentivity and unaccusativity to telicity	→ Agentivity and telicity are viewed as diametrically opposed.
(2) The weight of the two properties is asymmetrical	→ Telicity is the main factor.
(3) The position of the 0-point	→ It is difficult to understand if it is in the middle or at the bottom.
(4) Less attention to the most gradient verbs, and therefore to what determines the variation	→ The most gradient verbs, the ones in the middle, are the least taken into consideration.
(5) It is difficult to understand the lowest degree of telicity	→ It is difficult to understand if it is in the middle or at one end.
+(6) the exclusion of the reflexives	→ There is a gap in the analysis.

TABLE 1.11 – Table summarizing the limits of the ASH designed in Giancarli (2015: 82-86)

ASH posits that agentivity and telicity represent the extremes of a shared continuum: at one pole, telicity aligns with unaccusativity and verbs selecting ‘be’, while at the opposite end, agentivity aligns with unergativity and verbs selecting ‘have’.

Giancarli (2015: 82) raises two relevant questions: firstly, given that agentivity and telicity are not antonyms nor complementary terms, how can they occupy opposite ends of the same scale? Secondly, and in relation to the aforementioned idea, some verbs convey both telicity and agentivity—and so the question arises as to where these verbs should be positioned. For instance, Giancarli (2015: 82) mentions location verbs in French like *arriver* ‘arrive’ and *partir* ‘leave’.

A second issue is the asymmetry of the scale, with telicity exerting greater influence than agentivity. Telicity spans the entire scale, affecting all intransitive verbs, ranging “from a high degree of telicity at the top (the unaccusative end) to a low degree at the bottom (the unergative end)” (Giancarli 2015: 83). Verbs at either extreme are prototypical for one auxiliary, while those in the middle are paradoxically considered peripheral (Giancarli 2015: 81). In contrast, agentivity pertains only to verbs selecting ‘have’ at one end of the spectrum.

The third problem concerns the positioning of the zero point, namely the part of the scale where telicity and agentivity are at their lowest degree. As mentioned a few lines above, telicity

covers the whole scale and has its most downward influence at the bottom, the zone dedicated to ‘have’-selecting verbs. But agentivity only concerns the unergative part of the scale, so its zero point would be the middle of the scale.

The fourth point represents the heart of the question designed by a study on the auxiliary variation, namely the fact that “the most gradient verbs (those directly open to variation since they admit competition between the two auxiliaries and which, in Italian for example, are situated in the middle of the scale, see Legendre and Sorace 2003: 197 or Sorace 2004: 260)) are the least taken into account by the gradient representation developed by the ASH, since they are underspecified” (Giancarli 2015: 84-85). This is a crucial point as a study on the auxiliary variation should focus on the verbs most affected by this phenomenon, namely the ones in the middle.

Finally, the fifth point relates to the position of atelic verbs within the hierarchy. Considering that verbs characterized by graduality correlate least with unaccusativity and unergativity, they occupy a middle position. However, Giancarli (2015) highlights a problem: while atelic verbs are found in the middle of the hierarchy, telicity should theoretically cover the entire scale, suggesting that these verbs should be at the bottom. This discrepancy challenges the expected distribution of telicity across the hierarchy.

Moreover, when talking about auxiliary variation, we should not forget reflexives, which belong to the domain of the intransitive verbs too. Giancarli (2015: 86) claims that the possible reason why these verbs have not been taken into account is the absence of variation in the languages considered. But, even if it were, ASH would not be able to explain this phenomenon within reflexives as “telicity, the main factor used by the ASH with regard to intransitives, has already been tested for reflexives and cannot be sustained for the latter [...] Reflexives cannot belong to the domain of the ASH since it cannot account for them” (Giancarli 2015: 86). This could be a limit because, as the author of this work states, reflexives are equally concerned by auxiliary variation and are intransitive, necessitating an explanation that takes both into account.

Giancarli (2015: 83) also argues that ASH fails to accommodate substantial data in other Romance languages, viz. Corsican and French, where (lack of) telicity proves not to be relevant for verbs such as *rester* ‘stay’ or *demeurer* ‘remain’ that select ‘be’, despite their atelic nature—a paradox noted by Legendre and Sorace (see Giancarli 2015 for details).

Another issue that characterizes ASH is scalarity, which Giancarli illuminates with some Corsican examples. In Corsican, “three modal verbs and three aspectual ones [...] share the particularity of being transparent [with respect to auxiliary selection]: they can be preceded by either auxiliary” (Giancarli 2015: 87). What counts is the auxiliary selected by the infinitive

verb. But it happens that when a ‘be’ selecting verb is elided after one of these specific verbs, ‘have’ replaces ‘be’ as the default auxiliary, as in the following example:

- (19) De Soto era andatu al di là di ciò ch’iddu avaria pussutu.  
 De Soto be.IPFV.3SG go.to. PTCP beyond.of what that.he have.COND.3SG can.PTCP.

‘De Soto [...] had gone beyond what he could have.’

(Giancarli 2015: 87-88)

As we can see, “the ellipsis of the ‘be’ selecting verbs suppresses the requirement for the expected auxiliary (‘be’) and ‘have’ becomes possible. ‘Have’ is chosen, while a scalar representation [like ASH] would have predicted the selection of ‘be’ for these verbs” (Giancarli 2015: 89). Giancarli (2015: 89) emphasizes these ‘be’-verbs’ inability “to determine the choice of auxiliary” not only when they are omitted, but also when they are relocated in a different part of the sentence than the one expected.

The last issue concerns the cross-linguistic validity of the ASH: “to set out a wide variety of languages on a single scale, one must either take for granted in accordance with the Null hypothesis that all the compound forms (all perfects, all pluperfects, etc.) of the languages under consideration are equivalent to each other, or disregard the fact that they are not” (Giancarli 2015: 90).

Once again, the examples presented are from French and Corsican: while French perfect, the *passé composé*, can be used both with an anterior temporal reading and an aspectual one, Corsican *passatu quancianu* only presents the latter one, thus needing to use an aorist form without auxiliary when a temporal value is required. Therefore, “treating them identically nonetheless is one of the main reasons for the difficulties experienced by the ASH, since telicity, its primary factor on the gradient whatever the language, faces difficulty in French [...], is suited neither to German, nor Corsican [...] nor Arcadian [...], and is challenged by Italian and Dutch data” (Giancarli 2015: 92).

Giancarli (2015: 92) goes further by arguing that “any scalar representation of auxiliaries in a cross-linguistic perspective would meet difficulties” since the use of the same scale presupposes that the different languages are only different in quantitative terms. As a result, it seems that ASH may not be the best tool because, as a scalar representation, it is not adequate for describing such an unstable category like auxiliaries. As reported by Giancarli (2015: 92), McFadden and Alexiadou (2010) claim that “two languages will differ in their auxiliary selection at least in part because of the differences in the make-up of their perfects.”



correspond to Source and Goal. It categorizes verbs into three groups: process verbs without variation, process verbs with variation, and state verbs, as elaborated in Giancarli's detailed exposition (2015), and introduces a quantitative approach that accommodates the complexities of auxiliary selection and acknowledges the diversity across languages.

While Giancarli effectively highlights some limitations of the Auxiliary Selection Hierarchy (ASH), his theory does not fully account for all factors influencing auxiliary selection in intransitive verbs. Indeed, his theory is based on macroroles: thematic roles are emphasized, but telicity takes a secondary role.

### 1.7 CONCLUSION: SYNTHESIS AND PATHWAY TO METHODOLOGY

The analysis of auxiliary selection—the choice of either 'have' or 'be' for the formation of compound tenses in Italian and French—involves the intricate relationship between lexical semantics and syntax, which has been thoroughly examined through various theoretical frameworks.

The Unaccusative Hypothesis, originally proposed by Perlmutter (1979), provides a foundation for understanding the syntactic distinctions between unaccusative verbs, which tend to select 'be' in Italian and French, and unergative verbs, which select 'have'. Syntactic approaches have been further explored by Burzio (1986), Grimshaw (1990), and Levin & Rappaport Hovav (1995), highlighting the interplay between syntactic configurations and semantic roles.

Van Valin's semantic perspective adds depth by emphasizing the role of aspectual properties in auxiliary selection, while Sorace's Auxiliary Selection Hierarchy (ASH) introduces a nuanced gradient approach based on two parameters: telicity and agentivity.

Bentley and Eythórsson support ASH and further refine it by demonstrating that auxiliary selection patterns reflect underlying semantic distinctions associated with unaccusative verbs. Despite its comprehensive nature, the ASH has faced critiques regarding its handling of graduality and scalarity, as well as its applicability across languages.

Giancarli's thematic macro-roles theory offers alternative perspectives that address some of these limitations but does not provide a comprehensive explanation of all the factors influencing the choice of auxiliary, as telicity is relegated to a secondary position.

Considering the most important work that shaped the literature on auxiliary selection, there was a transition from a binary approach (unaccusativity correlates with 'be' and unergativity with 'have') to a more gradient approach. For example, Sorace (2000) demonstrates that unergative

verbs still correlate with ‘have’ and unaccusative with ‘be’ because of agentivity and telicity, respectively, and from one extreme to the other, there is a gradience of verbs that cannot be categorized in a binary way between verbs that select ‘have’ or ‘be’. Although this approach encompasses all the essential factors in auxiliary selection, evaluating the weight of each factor is still challenging.

Between the initial binary approach and Sorace’s gradience, other approaches have highlighted the importance of thematic roles and the aspectual dimension, even though in different ways (for example, Grimshaw 1990 and Van Valin 1990). The last two approaches analyzed in this chapter emphasize telicity (Bentley and Eythórsson 2003) and the thematic dimension (Giancarli 2015), demonstrating the complexity of auxiliary selection and the necessity of multifaceted analytical frameworks.

An approach that considers agentivity and telicity equally is necessary to understand their effective weight. As highlighted by Giancarli, attention to the most gradient verbs (namely, peripheral verbs) is essential to fully explain this phenomenon.

In this work, for example, we start by considering all the possible influencing factors, ultimately weighing each factor depending on the verb type. I anticipate here that peripheral verbs are less involved in telic contexts: this does not mean that telicity is unimportant in auxiliary selection, but rather that these types of verbs are less involved due to their own characteristics.

To achieve this, we have opted for a data-driven approach that considers both qualitative and quantitative corpus analysis. This approach helps understand the use and weight of each factor influencing auxiliary selection. Another question we aim to address is whether the differences between Italian and French are purely quantitative or scalar in nature (i.e., less often ‘be’, according to a scale), or if there are specific factors in French that are not at work in Italian.

This research project stems from a desire to understand the weight of various factors commonly considered influential in auxiliary selection. More specifically, my research tackles the verbs least addressed, viz. the peripheral verbs, in academic literature—those most subject to the selection of both auxiliaries—to understand the deepest possible level of factors influencing the selection of ‘have’ or ‘be’. Filling this gap is crucial for understanding the relationship between various factors that are normally considered important (e.g., agentivity and telicity among others).

Our findings show that the factors important to auxiliary selection differ in Italian and French. In Italian, as expected, telicity is not particularly influential for this class of verbs that can select both auxiliaries. The occurrences we analyzed show that the small differences leading to the

selection of ‘have’ or ‘be’ are semantic, especially distinguishing between subjects characterized by causality (internal cause, as we shall see) or agentivity.

For French, the division of verbs into different classes is necessary because it results in some classes where the selection of ‘have’ or ‘be’ depends on the construction (e.g., spatial motion constructions or verbs of appearance) and another that behaves like Italian (the change of state).

As we move into the next chapter, the methodology, we will outline the specific approaches and techniques employed to investigate these theoretical constructs. This methodological framework will provide the tools needed to empirically test and validate the hypotheses discussed, offering a robust foundation for analyzing auxiliary selection across different linguistic contexts.

## CHAPTER 2: METHODOLOGY

In research on auxiliary alternation, scholars often focus on introspection, linguistic intuition and acceptability judgments for establishing which verbs accept ‘have’, ‘be’, or both. This tendency is exemplified, for instance, by Sorace’s major study, developing the Auxiliary Selection Hierarchy hypothesis (ASH). The ASH outlines auxiliary selection in intransitive verbs by raising the question which intransitive verbs or types of intransitive verbs admit either ‘have’ or ‘be’ or both auxiliaries, and in which contexts, but without observing or quantifying the actual usage of auxiliaries based on empirical data.

Therefore, the aim of this study is to reassess Sorace’s hierarchy (2000) on the basis of a quantitative estimate of the use of ‘have’ and ‘be’ with certain specific verbs, obtained through corpus analysis, rather than relying on acceptability judgments on the possible combination of these auxiliaries with specific verbs. My study will focus on a particular and understudied area of the hierarchy, namely the so-called peripheral verbs (i.e. intransitive verbs allowing for variation) rather than encompassing all intransitives. It is expected that the study of these verbs allowing to different degrees both auxiliaries will contribute to a comprehensive understanding of auxiliary selection.

A qualitative corpus analysis will be employed to elucidate both the selection of a particular auxiliary and the factors influencing this selection. Moreover, a statistical analysis will be performed on the corpus to identify the factors that consistently influence auxiliary selection.

In the first part of this chapter (2.1), I will focus on the corpus data, motivating the chosen corpora in section 2.1.1, and explaining the methods for the qualitative corpus analysis in section 2.1.2. Section 2.1.3 will detail the parameters systematically used in the qualitative corpus analysis, whose impact on auxiliary selection will be measured. The second part of this chapter (2.2) will focus on the quantitative corpus analysis, presenting the statistical methods used.

## 2.1 CORPUS DATA

### 2.1.1 CHOICE OF THE CORPORA

Corpus analysis is a pivotal aspect of this study, since it fills a gap in the state of the art. Given its comparative nature, examining auxiliary selection in verbs exhibiting variation across Italian and French necessitated the selection of appropriate corpora for analysis. The chosen corpora, ItTenTen (16) and FrTenTen (17), are both members of the TenTen family and are accessible via SketchEngine<sup>6</sup> (Kilgarriff et al. 2014). SketchEngine proves instrumental in linguistic analysis, leveraging algorithms to sift through vast amounts of authentic text, enabling the identification of typical language patterns as well as emerging usages.

ItTenTen, also known as the *Corpus of the Italian Web*<sup>7</sup> comprises a substantial compilation of web texts. The 2016 version containing 4.9 billion words, was selected for this study due to its extensive and diverse content, including materials from online newspapers, ensuring a blend of formal and informal language. Lemmatization using the Tree Tagger tool, based on Marco Baroni's parameters<sup>8</sup>, enhances the corpus's utility for detailed linguistic analysis.

FrTenTen, or the *Corpus of the French Web*<sup>9</sup>, mirrors ItTenTen in its internet-based composition. The version used in this study, FrTenTen17, comprises 5.7 billion words and offers similar advantages to its Italian counterpart. Despite differences in Part of Speech (POS) tagging compared to ItTenTen16, as detailed in section 2.1.1, FrTenTen17 remains a valuable resource for comparative analysis.

The decision to focus on internet-based corpora stems from two primary considerations. Firstly, the vast volume of data provided by these corpora facilitates in-depth research. Secondly, a balanced analysis necessitates consideration of both formal and informal registers. While initially considering the use of exclusively journalistic corpora, challenges arose due to register differences between Italian and French newspapers: whereas Italian newspapers adopt a more informal register, French journalists tend to adopt a more controlled and normative register (Karssenber 2018: 236). Another option considered were spoken language corpora. Again, sourcing suitable and comparable datasets for the two languages proved challenging. Therefore, web corpora emerged as the most viable option.

---

<sup>6</sup> <https://www.sketchengine.eu>).

<sup>7</sup> <https://www.sketchengine.eu/ittenten-italian-corpus/>.

<sup>8</sup> <https://www.sketchengine.eu/italian-treetagger-part-of-speech-tagset/>

<sup>9</sup> <https://www.sketchengine.eu/frtnten-french-corpus/>.

## 2.1.2 METHODS FOR QUALITATIVE CORPUS ANALYSIS

The selected corpora, ItTenTen (16) and FrTenTen (17), formed the foundation for both qualitative and subsequent quantitative analyses, serving as the primary dataset. For qualitative analysis, the Corpus Query Language (CQL) tool within Sketch Engine was employed, facilitating advanced search capabilities for identifying verbs allowing for two different auxiliary verbs.

The coding process varied slightly between languages due to differences in Part of Speech (POS) tagging methods. As mentioned previously, ItTenTen (16) utilized the Tree Tagger tool, developed by Helmut Schmid at the Institute for Computational Linguistics of the University of Stuttgart, employing Marco Baroni's parameter file<sup>10</sup>. In contrast, FrTenTen (17) was annotated using the FreeLing tool<sup>11</sup>.

PART OF SPEECH (POS)	ITENTEN (16) TAGSETS	FRENTEN (17) TAGSETS
Noun	NOUN	N.*
Pronoun	PRO	P.*
Past Participle	VER: ppast	V.P.*

TABLE 2.1 - Part-of-Speech different tagsets in ItTenTen (16) and FrTenTen (17)

For instance, nouns are explicitly labeled as 'NOUN' in ItTenTen, while in FrTenTen, they are abbreviated as 'N.\*'. Similarly, pronouns are tagged as 'PRO' in ItTenTen and 'P.\*' in FrTenTen, and past participles are marked as 'VER: ppast' in ItTenTen and 'V.P.\*' in FrTenTen.

These tags differences are crucial for conducting syntactic structure searches via CQL to identify perfective constructions involving two auxiliaries. The difference in POS tagging (Free Ling Tool or Marco Baroni's parameters) will not have an impact on simple sentences, such as those described in the queries involving nouns, auxiliaries, and past participles.

Hence, the following CQL commands were employed using these tagsets:

- (1) ITALIAN: [tag= "NOUN|PRO: pers"] [lemma= "avere|essere"] [tag= "VER: ppast"]
- (2) FRENCH: [tag= "N.\*|P.\*"] [lemma= "avoir|être"] [tag= "V.P.\*"]

<sup>10</sup> <https://www.sketchengine.eu/italian-treetagger-part-of-speech-tagset/>

<sup>11</sup> <https://www.sketchengine.eu/frtnten-french-corpus/>

By employing this search syntax, it became possible to identify sentences within the corpora containing a subject—either a personal pronoun or a noun—an auxiliary ‘have’ or ‘be’ in all their forms, and a past participle. Despite Italian being a pro-drop language, allowing for subject-less phrases, we included this criterion to streamline the search in a vast corpus and maintain consistency with French, where subjects are obligatory.

Additionally, the rule that no noun should immediately follow the past participle within five words was established to facilitate the exclusion of transitive constructions (although complete exclusion proved impossible). For example, with the part-of-speech filter ensuring that no noun appears within the five words following the past participle, the CQL should be able to filter out transitive constructions such as *Ho cambiato casa* (‘I changed flat’) or *Ho cambiato volentieri e diverse volte casa* (‘I changed flat willingly and several times’) for verbs that can be used both transitively and intransitively. This approach maximizes the likelihood of capturing the intransitive verb construction, which interests us in the context of the present study.

The utilization of CQL facilitated the identification of new verbs warranting investigation to ascertain whether they exhibited variation. Consequently, we compiled a dataset deemed sufficient to fulfill two objectives: 1) ensuring a representative corpus for analysis and 2) enabling a robust qualitative analysis while retaining the flexibility to incorporate both qualitative and quantitative approaches.

The initial step involved compiling a list of potential verbs demonstrating variation. The table in section 2.2 shows the list of verbs generated by analyzing the first 2000 occurrences in ItTenTen16, utilizing the Italian CQL query mentioned in (1), namely [tag= “NOUN|PRO: pers”] [lemma= “avere|essere”] [tag= “VER: ppast”].

<b>ITTEN (16): LIST OF VERBS</b>	
1.	<i>Cambiare</i> ‘change’
2.	<i>Cedere</i> ‘yield’
3.	<i>Suonare</i> ‘ring’
4.	<i>Contare</i> ‘count’ ‘matter’
5.	<i>Pesare</i> ‘weigh’ ‘matter’
6.	<i>Prevalere</i> ‘prevail’
7.	<i>Fallire</i> ‘fail’
8.	<i>Iniziare</i> ‘start’
9.	<i>Cominciare</i> ‘begin’
10.	<i>Procedere</i> ‘go on’
11.	<i>Continuare</i> ‘continue’
12.	<i>Proseguire</i> ‘go on’
13.	<i>Finire</i> ‘finish’

TABLE 2.2 - List of Verbs with Auxiliary Selection emerged in ItTenTen (16)

Table 2.3 shows the list of verbs generated by analyzing the first 2000 occurrences in FrTenTen (17) utilizing the French CQL mentioned in (2), namely [tag= “N.\*|P.\*”] [lemma= “avoir|être”] [tag= “V.P.\*”].

<b>FRTEN (17): LIST OF VERBS</b>	
1.	<i>Passer</i> ‘go on’
2.	<i>Monter</i> ‘go up’ ‘increase’
3.	<i>Baisser</i> ‘go down’ ‘decrease’
4.	<i>Descendre</i> ‘go down’
5.	<i>Changer</i> ‘change’
6.	<i>Mûrir</i> ‘mature’
7.	<i>Diminuer</i> ‘go down’ ‘decrease’
8.	<i>Exploser</i> ‘explode’
9.	<i>Paraître</i> ‘seem’ ‘appear’
10.	<i>Apparaître</i> ‘seem’ ‘appear’

TABLE 2.3 - List of Verbs with Auxiliary Selection emerged in FrTenTen (17)

In Tables 2.2 and 2.3, only one verb (e.g., ‘change’) is shared between Italian and French verbs. This result already shows an important difference between Italian and French with respect to auxiliary selection: despite their common Latin ancestor, the auxiliary selection of the Italian and French verbs is different.

As a matter of fact, the lists of Italian and French verbs differ significantly in both number and type, with *cambiare* (‘change’) as the sole verb common to both. This difference results from a combination of rigorous analysis and inherent linguistic variation, as will become evident. The initial analysis involved 2000 sentences for both Italian and French, using the CQL query mentioned earlier. Verbs exhibiting auxiliary alternation were further analyzed within a smaller sample 100 sentences, selected using the random sampling function provided by SketchEngine. As noted by SketchEngine, “a random sample is used to reduce the number of concordance lines while preserving the representativeness of the sample.” For certain verbs, the query did not yield a sufficient number of occurrences, resulting in samples that may contain fewer than 100 sentences.

If one auxiliary predominates within these sampled sentences, it is a strong indication that the verb in question does not exhibit alternation, particularly after exclusion of instances involving passive or transitive constructions. The final list comprises thirteen verbs that were identified as occurring with both ‘have’ and ‘be’ in seemingly similar contexts, which were deemed particularly relevant for further analysis.

Since the Italian verbs were analyzed prior to the French ones, an attempt was made to examine the same verbs in French. However, the analysis reveals significant differences between the two languages, and several verbs that allow auxiliary alternation in Italian do not exhibit this behavior in French. Not only do the verbs allowing auxiliary alternation differ, but also factors influencing auxiliary selection vary between Italian and French and the option of using the same list for both languages would therefore not align with my research project which aims to identify the parameters that determine auxiliary selection for verbs with auxiliary alternation. As a result, the Italian and French verb lists are not identical.

Nevertheless, the same methodology was applied rigorously and consistently to both languages. While the resulting lists differ, this ensures that the methods were unbiased and systematically implemented. To create a coherent corpus that accurately reflects the alternation of ‘have’ and ‘be’—without introducing selection bias—it was essential to use a random corpus and apply the CQLs as specified. This approach ensured the identification of verbs that reflect actual usage patterns rather than conforming to preconceived models and prevented that hypotheses valid for one language be projected onto the analysis of the other.

The second stage of the analysis consisted of collecting occurrences for each selected verb one at a time. The same CQL methodology was employed, with the participle of the verb under investigation being lexically instantiated by each of the verbs of both lists (Tables 2.2 and 2.3). For each verb of both lists, queries were made with the auxiliary ‘have’ and with the verb ‘be’, resulting in 50 sentences with ‘have’ and 50 sentences with ‘be’ for each verb in both languages, whenever possible. To exclude transitive sentences, we maintained the condition: no noun after the participle within the first 5 words.

For example, considering the verb *cambiare* in Italian, the CQL would be as follows: [tag=“NOUN|PRO:pers”] [lemma=“avere”] [tag=“cambiato|cambiata|cambiati|cambiate”] and [tag=“NOUN|PRO:pers”] [lemma=“essere”] [tag=“cambiato|cambiata|cambiati|cambiate”], with the Part of Speech Filter excluding selected nouns thereafter. Additionally, the ‘random sample’ option was incorporated to ensure that the 100 occurrences were as representative of the corpus as possible. The selection of the number of occurrences was driven by the need to maintain a corpus that was both sufficiently large and analytically traceable.

During the detailed analysis of the verbs, it became apparent that some Italian verbs did not provide enough intransitive occurrences with both ‘have’ and ‘be’. As a result, only 1183 occurrences were suitable for the study. Furthermore, four of these verbs exhibited additional usages, which were subsequently analyzed as semi-auxiliaries. Combining the 1183 occurrences with the 400 from the first analysis, the total occurrences analyzed for Italian verbs amounted to 1583.

Regarding French verbs, nearly all verbs yielded 100 occurrences, with the exception of one verb (*mûrir*), resulting in a total of 990 occurrences. Altogether, 2573 occurrences comprising both French and Italian verbs were manually analyzed. For these verbs, sixteen parameters were devised to ascertain the overall influence of all factors.

### 2.1.3 PARAMETERS FOR QUALITATIVE CORPUS ANALYSIS

The second part of the qualitative analysis section consisted in the investigation of various parameters that could influence the selection of the auxiliary. Specifically, an assessment was conducted to determine the parameter characterizing each sentence, categorized into two types: semantic and syntactic. The latter category was further subdivided into adverbials and other syntactic parameters, including the absence of adverbials, subject-past participle agreement, and the presence of an implied direct object.

The different parameters are represented in Table 2.4.

SYNTACTIC PARAMETERS (OTHERS)	SYNTACTIC PARAMETERS (ADVERBIALS)	SEMANTIC PARAMETERS
[± Zero] (Free-adverbials context) [± Agreement subject – past participle]	[± Adverbial of manner]	Subject [± Human]
	[± Adverbial of time]	Subject [± Agentive]
[± Direct Object implied]	[± Adverbial of time + duration]	Subject [± Animate]
	[± Locative Adverbial] Static	Subject [± Internal cause]
	[± Locative Adverbial] [-endpoint]	
	[± Locative Adverbial] [+endpoint]	
	[± Aspectual Adverbial] Telic	
	[± Aspectual Adverbial] atelic	
	[± Aspectual Adverbial] unmarked	
	[± Adverbial of quantity]	
	[± Argument Adverbial]	

TABLE 2.4 – Schema of the parameters considered in the qualitative analysis.

*Syntactic parameters (others)* These syntactic factors include the free-adverbial context, agreement between subject and past participle, and the presence of an implied direct object. Regarding agreement, the analysis aimed to investigate whether sentences with feminine subjects exhibit different auxiliary selection patterns compared to those with masculine subjects. This investigation is pertinent because more attention is typically paid to the auxiliary when the past participle agrees with the subject. In cases where the subject is masculine in Italian, the past participle remains in its neutral form. In French, the singular and plural forms are homophonous. However, this parameter revealed no influence over the distribution of the auxiliaries. The free-adverbial context and the agreement between subject and past participle proved not to be influential. However, it was important to consider the hypothesis of a potential impact before excluding it.

Finally, the analysis considered instances where sentences could contain an implied direct object (DO). Despite efforts to exclude transitive phrases, some instances still occurred, and interpretation could be ambiguous, particularly when there is no expressed object or when there is a potential ambiguity between an object complement or an adverb, as is the case in (1).

(1) *Luigi ha cambiato molto* ‘Luigi has changed a lot’

The ambiguity of this sentence hinges on the word *molto*, which in Italian can function as both an adverb and an adjective. In the latter case, the sentence would imply that ‘Luigi has made numerous changes’, in the sense of changing many furnishings in the house. At the same time, the verb ‘change’ affecting the subject, such as in *dopo il matrimonio, Luigi è/ha cambiato molto* (‘after marriage, Louis has/has changed a lot’). While the transitive interpretation might seem more plausible with ‘have’ (as the intransitive ‘have’ may appear strange), we cannot completely rule out the other option without considering various data providing different contexts.

The possibility of ambiguity may be heightened by the fact that many of these verbs can be used both transitively and intransitively. Consequently, we opted not to exclude these data entirely. Analyzing them alongside actual transitive uses enables us to better understand the factors influencing selection, effectively serving as a form of control test. This approach allows us to delineate the thin line that distinguishes transitive from intransitive interpretation.

***Syntactic parameters (adverbials)*** Various types of adverbials have been examined to encompass a broad spectrum of contexts. The conceptual framework for these adverbials has been shaped through the consultation of works by Blumenthal (1990) and Carlier & Sarda (2010). These include manner adverbials, quantity adverbials, and argument adverbials. Additionally, aspectual adverbials conveying telic aspect, atelic aspect, or unmarked with respect to telicity and denoting progressive duration were considered. Time adverbials were categorized into simple time adverbials and those denoting duration. Lastly, locative adverbials were classified into those indicating a static place, those indicating a direction with an endpoint, and those indicating a direction without an endpoint.

***Semantic parameters and the introduction of Internal Cause*** With respect to semantic factors, we examined the subject’s semantics, which led to the delineation of traits such as [ $\pm$ HUMAN], [ $\pm$ ANIMATE], [ $\pm$ AGENTIVITY], and [ $\pm$ INTERNAL CAUSE]. It is important to distinguish between animacy and agentivity, as well as between agentivity and internal cause. While Sorace (2000) sometimes treats animacy and agentivity as overlapping concepts, my analysis sought to avoid conflating the two. Although animacy and agentivity may be related, not all animate subjects are necessarily agentive, whereas it is impossible for inanimate subjects to exhibit agency.

In my analysis, a further distinction was introduced between agentivity and internal cause (see Cruse 1973; Demirdache & Martin 2015; Finocchiaro 2002; Fritz-Huechante et al. 2020; Martin 2020 for more details about agentivity and causation). To deepen this exploration, the concept of non-volitional internal cause from Mateu (2009) was incorporated, explaining why some intransitive verbs select ‘have’ rather than ‘be’. This concept was extended beyond verbs typically selecting ‘have’ to include those demonstrating split intransitivity. Integrating Mateu’s concept with Reinhart’s (2000, 2002) hypothesis, which defines agentivity by the presence of ‘c’ (cause change) and ‘m’ (mental state), allowed internal cause to be interpreted as a subset of agentivity, namely the presence of ‘c’ without ‘m’.

Internal cause is the property of the subject being both the cause and the patient of the action. For instance, in (2):

(2) *Marie a changé profondément.*

‘Marie has changed profoundly’.

In the French example in (2), Marie can be the agent of her change, by wanting to improve herself, and at the same time the patient of it, as she might change involuntarily after certain experiences. This dual role fits my definition of an internal cause. This concept can be due to the semantics of the verb, being causative without being voluntary, or to the features of the subject, for instance in the case of nouns evoking activity when there is a person behind them causing the action despite the instruments being the subjects. For example (3):

(3) *Numerosi studi hanno continuato nel 1990 e anche dopo la morte del dottor Atkins.*

‘Numerous studies continued in the 1990s and even after Dr. Atkins’ death’.

In sentence (3), *studi* ‘studies’ are the subject, and the actions continue due to the previous setup by the doctor, indicating a shifted intentionality. This aspect will be better demonstrated in Chapter 3, dedicated to the corpus analysis of Italian.

The notion of internal cause in my analysis stems from challenges in identifying what constitutes agentivity, as already seen with Mateu (2009) and Reinhart (2000, 2002). It is important to differentiate internal cause from agentivity: internal cause represents a subset of causality characterized by intrinsic causality lacking intention or control.

Therefore, agentive subjects must possess internal cause, along with the trait [+control] and/or intentionality (Figure 2.1).

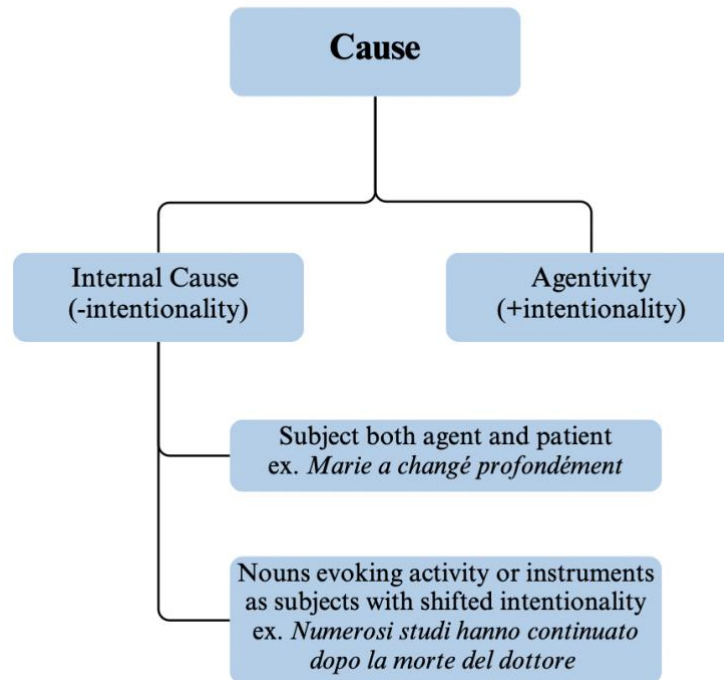


FIGURE 2.1 – Distinction between internal cause and agentivity.

Figure 2.1 visualizes the distinction between internal cause and agentivity. According to the diagram, the category of cause can branch into ‘internal cause’ and ‘agentivity’.

**Internal cause** characterizes subjects conceived as a cause and lacking intentionality. It can occur in sentences where the subject is both agent and patient, as in *Marie a change profondément* (example 2). It is also characteristic of subjects presenting shifted intentionality. An example is *Numerosi studi hanno continuato nel 1990 e anche dopo la morte del dottor Atkins* (example 3), where the studies act due to the prior intentionality of a person.

**Agentivity** combines the features of causality and intentionality. This distinction allows us to characterize for instance inanimate subjects that lack volition but still exert causative influence in some way. The broader applicability of the concept of internal cause, in comparison with agentivity proved to be necessary to formulate a comprehensive analysis of auxiliary selection.

The introduction of the concept aligns well with the existing literature on causativity and agentivity. DeLancey (1984) outlines the prototype of agentivity as comprising two key traits: volition and causation. According to DeLancey (1984), an entity qualifies as an agent when it

alone can be identified as the cause of an event, regardless of whether the action is voluntary or not. Similarly, Schlesinger (2006) proposes three agentive traits: cause, control, and change. He suggests that the trait [+cause] corresponds to [+causation], and [+control] corresponds to [+volition] and argues that the presence of any one of these traits increases the likelihood of classifying a subject as an agent, even if not all three traits are present. The absence of traits such as [+volition] or [+control] often leads to ambiguity regarding agentivity. However, by applying the concept of internal cause, and linking agentivity to the presence of both [+cause] and [+volition], this ambiguity can potentially be resolved.

**Parameters for semi-auxiliary verbs** As mentioned in 2.1, certain Italian verbs showed a distinct behavior when used as full verbs in comparison with their use as semi-auxiliaries. Specifically, four verbs—namely *iniziare*, *cominciare*, *continuare* and *finire*—exhibited such behavior. Semi-auxiliaries (e.g., *Matilda ha continuato a lavorare duro* ‘Matilda continued to work hard’), as will be elaborated upon in the dedicated chapter, possess attributes of both main verbs and auxiliary verbs, setting them apart from full auxiliaries and full verbs.

Semi-auxiliaries typically combine with a preposition followed by an infinitive, which may be either transitive or intransitive. Consequently, the semantic parameters devised for semi-auxiliaries align with those for full verbs, while an additional consideration for semi-auxiliaries pertains to the type of infinitive, as illustrated in table 2.5.

SEMANTIC PARAMETERS		SYNTACTIC PARAMETERS	
<i>Subject</i> [± Human]	<i>Subject</i> [± Animate]	<i>V AUX + V INF Trans</i>	<i>V AUX + V INF</i> <i>Intransitive Aux. ‘have’</i>
<i>Subject</i> [± Agentive]	<i>Subject</i> [± Internal cause]	<i>V AUX + V INF</i> <i>Intransitive Aux. ‘be’</i>	<i>V AUX + V INF</i> <i>Intransitive Aux. ‘have’/</i> <i>‘be’</i>

TABLE 2.5- Schema of the parameters considered in the qualitative analysis.

**Syntactic parameters (infinitives)** As depicted in table 2.5, the types of infinitives following the prepositions are categorized into four types according to their syntactic properties: transitive verbs, intransitive verbs selecting ‘have’ (e.g., *lavorare* ‘work’), intransitive verbs selecting ‘be’ (e.g., *arrivare* ‘come’), and finally, intransitive verbs compatible with both auxiliaries,

such as verbs denoting atmospheric events (e.g., *piovere* ‘rain’). The analysis will reveal that the type of verb in the infinitive form serves as one of the primary factors influencing the distribution of auxiliaries.

***Verbs, Verbal Constructions and Auxiliary Alternation*** Since Perlmutter’s (1978) Unaccusative Hypothesis, it has been customary to consider the choice between the auxiliaries ‘be’ and ‘have’ as one of the criteria for dividing intransitive verbs into two classes: unergative and unaccusative. According to this hypothesis, ‘have’ as a perfective auxiliary indicates that the single argument of the intransitive verb is syntactically equivalent to the subject of a transitive verb and is semantically an agent. In contrast, when the auxiliary ‘be’ is used, the single argument is conceived as being syntactically equivalent to the object of a transitive verb and having patient-like features. The present corpus study focusing on peripheral verbs—compatible with both ‘be’ and ‘have’—which take into account various usage patterns, will show that auxiliary selection is not strictly constrained by the semantics of the verb verbs but that verbs may select a different auxiliary according to the construction in which they are embedded. An example is provided by the verb *descendre* ‘go down’ used in a motion verb construction in (4) and in a quantitative evolution construction in (5)

(4) *Le pilote était descendu pour être au-dessous des nuages [...]*

(5) *Son taux d'oxygène dans le sang est descendu à 30%*

(FrTenTen (17))

In examples (4) and (5), the auxiliary verb depicted is ‘be’. However, this study will demonstrate that the type of construction (e.g., spatial motion vs quantitative) and the intrinsic features of the subject—such as [ $\pm$  AGENTIVE], [ $\pm$ INTERNAL CAUSE], and [ $\pm$ HUMAN]—can also influence the choice and frequency of the auxiliary verb. These factors determine whether ‘have’ or ‘be’ is used. For instance, locative constructions will demonstrate to present a preference for ‘be’, but in the locative construction *J’ai descendu dans mon jardin, pour y cueillir du romarin* (a popular child song, with an archaic use of ‘have’ (FrTenTen (17))), the selection of ‘have’ may be influenced by the agentive and human properties of the subject.

However, it may happen that the subject is ambiguous with respect to these features. A case in point is provided in the following Italian example:

(6) *woooooow flashgames ha cambiato .... era da un po' che non venivo piu... speriamo che anche i giochi siano migliorati*<sup>12</sup>

The auxiliary expected for this sentence in Italian would traditionally have been ‘be’. The presence of ‘have’ or the alternation between ‘have’ and ‘be’ in cases like (6) could be interpreted as an avant-garde occurrence of a diachronic tendency, wherein the morphosyntax of ‘have’ gradually permeates that of ‘be’, without any consistent accompanying semantic change. This historical tendency is clearly illustrated in Rosemeyer (2014: 4):

“[...] it has been hypothesized that gradient auxiliary selection in synchronic stages of Old and Early Modern Spanish is the result of a gradual process by which *haber* expanded into usage contexts previously reserved for *ser*.”

The broader tendency among European languages to favor ‘have’-selection has resulted in many languages abandoning ‘be’-selection altogether (Rosemeyer 2014: 32). However, in French and Italian, the division between ‘have’ and ‘be’ persists. According to Pountain (1985), this resilience can be traced back to Latin. The Latin verbs *tenere* and *stare* did not play a significant enough role to replace the functions of ‘have’ and ‘be’, respectively. Consequently, as Rosemeyer (2014:35) notes, “the use of ‘have’ was not strengthened, and the use of ‘be’ not weakened.”

However, considering Italian’s resistance to auxiliary change and analyzing the sentence further, the subject in (6) could be interpreted as internal cause. This is particularly plausible if we imagine an implicit action behind it, suggesting the involvement of an external agent effecting the change. This results in an ambiguity between interpretations such as “flashgames changed” and “flashgames changed its interface/system/etc.”

This ambiguity only arises when the auxiliary is ‘have’. Since the parameter to be predicted is the choice of the auxiliary, it may seem a methodological error of circular reasoning to use the presence of the auxiliary *avere* ‘have’ to interpret the subject *flashgames* as an internal cause, because this interpretation will then appear in the statistics for the parameters conditioning the choice of the auxiliary.

---

<sup>12</sup>English translation: “Wow flashgames has changed...it’s been a while since I’ve been here...hopefully the games have improved as well.”

I will present two arguments to justify my methodology. First, my hypotheses are not inferred from single examples but from robust statistical trends, and the few instances where the subject's properties are undecidable do not substantially alter these statistical trends. Second, it is important to consider that the lexicon is intrinsically flexible in terms of its interpretation, and in cases of discordance between grammar and the lexicon, the lexicon tends to be reinterpreted. An example from a different domain, namely nominal determination, will help illustrate my point.

The distinction between count nouns and mass nouns may initially appear as an intrinsic feature of nouns. In French, there is a clear indicator: count nouns combine with the singular indefinite article (e.g., *une pomme* 'an apple'), whereas mass nouns combine with a so-called 'partitive' article (e.g., *du sang* '(some) blood'). However, when Zola (*Au Bonheur des Dames*, 1883), writes *Sous la grâce même de sa galanterie, Mouret laissait ainsi passer la brutalité d'un juif vendant de la femme* 'Under the very grace of his gallantry, Mouret thus allowed the brutality of a Jew selling women', the noun *femme*, although intrinsically a count noun referring to individual human beings—when combined with the partitive article—yields a deindividuated view of women. This example demonstrates that, in cases of conflict between grammar (the partitive article) and lexicon (the noun *femme*), it is the grammatical construction that determines the interpretation.

In contrast to the traditional view that lexical units (whether verbs or nouns) determine possible constructions, the construction grammar perspective, as developed over the past three decades (Goldberg 1995), offers a more nuanced understanding. Goldberg (1995:189) explains:

“On a constructional approach to argument structure, systematic differences in meaning between the same verb stem in different constructions are attributed directly to the particular constructions.”

Since the auxiliary selection is the target variable, it is necessary to establish parameters that condition the selection of the auxiliary, such as the semantic features of the subject, independently of the auxiliary choice. However, applying the construction grammar framework to the alternation between 'have' and 'be' as auxiliaries, it becomes evident that there comes a point where, for isolated instances presenting ambiguity, we must acknowledge that the interpretation of the lexical content conveyed by the subject depends on the construction used and more specifically on the choice of 'have' as an auxiliary. Therefore, in ambiguous cases, such as the one presented in (6), it is hypothesized that the use of 'have' within a specific construction can favor the interpretation of the subject as an internal cause.

This perspective aligns with the construction grammar approach, which posits that grammatical constructions themselves carry meaning independently of the lexical units in a sentence. Therefore, the syntactic patterns associated with simple sentences are imbued with meaning, and the constructions themselves play a central role in the relation between form and meaning.

## 2.2 FROM QUALITATIVE TO QUANTITATIVE CORPUS ANALYSIS: STATISTICAL METHODS

The necessity of statistical analysis applied to the data analyzed with qualitative corpus analysis lies in its capacity to unveil underlying patterns and regularities within the data that may not be readily discernible through qualitative examination alone. While qualitative corpus analysis furnishes insights into language patterns and allows to formulate hypotheses with respect to auxiliary selection and variation, statistical analysis provides a quantitative, rigorous and reproducible validation method for testing these hypotheses.

Two statistical methods were employed on the sentences previously analyzed during the corpus analysis and the semantic and syntactic parameters described earlier (table 2.4): CHAID (Chi-square Automatic Interaction Detector, Kass 1980) and Random Forest (Breiman 2001; Liaw & Wiener 2002). Further explanation of these methods will be provided in the dedicated chapter.

**CHAID** was conducted using SPSS version 26 and STATA software version 15. This nonparametric test provides a method for predicting the auxiliary (target variable) based on the values of predictor variables, which include the sixteen parameters outlined in section 2.1.2. This is achieved by constructing a decision tree. Essentially, CHAID divides the data into mutually exclusive and exhaustive subsets that most effectively characterize the dependent variable (Kass 1980: 119).

**RANDOM FOREST** analysis was performed using the R-package “caret” version 6.0 (Kuhn 2008). Random Forest assesses the importance of various predictors in determining the target variable (auxiliary), shedding light on which variables exert the greatest influence on classification. This statistical method offers an additional perspective on the importance of each predictor in the conducted analysis, by providing a ranking of all the predictors in an ascending order.

The integration of these statistical techniques is considered crucial for gaining a comprehensive understanding of auxiliary selection. These methods empower us to either

confirm or question hypotheses formulated at the qualitative level, thereby uncovering insights that might have been less discernible if solely relying on qualitative factor analysis.

### 2.3 THE FRESH INSIGHTS FROM THIS ANALYSIS

This thesis is built upon a robust foundation of corpora analysis, integrating both qualitative and quantitative methodologies to unravel the underlying mechanisms behind various phenomena. By combining qualitative corpus analysis with statistical techniques, this study aims to elucidate and validate these mechanisms.

One significant aspect of this analysis is the possibility to reinterpret peripheral verbs in light of internal cause. Rather than solely focusing on agentivity, this thesis delves into how the factor of internal cause correlates with the use of 'have'. This exploration aims to shed light on the ambiguity surrounding these peripheral verbs, bridging the gap between the absence of agentivity and causative behavior.

Additionally, the specific behavior of certain verbs as semi-auxiliaries with respect to auxiliary selection in Italian will be emphasized. While this topic has been explored in existing literature under the name of 'Restructuring' (Cinque 2004; Rizzi 1982; Rizzi 1976a), this analysis will highlight the role of infinitives in conjunction with semantic parameters.

Lastly, the comparison between French and Italian will uncover how the distribution of the auxiliary reflects distinct linguistic requirements in each language and does not correspond to a simple difference in cut-off point on the same scale of the auxiliary selection hierarchy, as has been assumed in almost all recent studies on the subject. Despite the differences, some similarities in auxiliary usage can also be identified.

Having established a comprehensive methodological framework through both theoretical and practical analyses, we are now prepared to extend our investigation into the corpus analysis phase. The detailed examination of Italian and French verb usage will leverage the insights gained so far to further explore the factors influencing auxiliary selection.

By employing both qualitative and quantitative methods, we aim to provide a deeper understanding of the interaction between internal cause, agentivity, and auxiliary usage in real-world language data.



## CHAPTER 3: CORPUS ANALYSIS OF ITALIAN VERBS

This chapter presents the qualitative corpus analysis of parameters and results concerning Italian verbs. The analysis primarily involves 13 full verbs and 4 semi-auxiliaries of aspect.

The selection of verbs followed a two-step process. Initially, a specific Corpus Query Language (CQL) was used to identify verbs associated with both ‘have’ and ‘be’ in the ItTenTen (16) corpus. The first CQL employed for this purpose was described in Chapter 2, namely [tag= “NOUN|PRO:pers”] [lemma= “avere|essere”] [tag= “VER: ppast”].

The initial analysis of 2000 occurrences revealed that thirteen verbs exhibited variation: *cambiare* ‘change’, *cedere* ‘yield’, *suonare* ‘ring’, *contare* ‘matter’, *pesare* ‘matter’, *prevalere* ‘prevail’, *fallire* ‘fail’, *iniziare* ‘start’, *cominciare* ‘begin’, *procedere* ‘go on’, *continuare* ‘continue’, *proseguire* ‘go on’, and *finire* ‘finish’.

These 13 identified verbs can be categorized into different groups:

- (i) Change-of-state verbs (Levin 1993: 244): *cambiare*
- (ii) Verbs of sound emission (Levin 1993: 234-235): *suonare*
- (iii) Emotional state verbs (‘Psych-verbs’ in Levin 1993: 188): *contare, pesare*
- (iv) Verbs of existence (Levin 1993: 249): *prevalere*
- (v) Verbs of failure: *fallire*
- (vi) Verbs indicating a stage of a dynamic progress (start, progression, continuation, or completion): *iniziare, cominciare, procedere, continuare, proseguire, finire*

Category (vi) includes some verbs identified as ‘Aspectual verbs’ in Levin (1993: 274). However, this designation does not fully capture the verbs in question, but rather their use when employed as semi-auxiliaries. Therefore, a distinction is made in this study between the use of these verbs as fully lexical verbs, indicating a stage of a dynamic process, and their use as aspectual verbs (in section 3.9).

The analysis consisted of two searches: a first one involving a CQL query employed for each full verb, with (1) and (2) representing an example with the verb *cambiare*:

- (1) [tag= “NOUN|PRO:pers”] [lemma= “avere”] [word= “cambiato|cambiata|cambiati|cambiate”]

(2) [tag= "NOUN|PRO:pers"] [lemma= "essere"] [word= "cambiato|cambiata|cambiati|cambiate"]

A second search was devised for semi-auxiliary verbs, represented by (3) with the verb *finire* as an example:

(3) [tag= "NOUN|PRO: pers"] [lemma= "avere|essere"] [word= "finito|finita|finiti|finite"] [tag= "PRE"] [tag= "VER:infi"].

The data were then selected using Sketch Engine's 'random sample' function to obtain 50 sentences with 'have' and 50 with 'be', when possible, to ensure a limited but representative sample from the corpus. The sample size was adjusted when necessary due to a lack of suitable data in the initial 50. Not all verbs yielded 100 occurrences; hence, there were 1183 for simple full verbs and 400 for semi-auxiliaries, totaling 1583 sentences.

This amount of data was chosen to ensure sufficient material for analysis without overwhelming the study. Moreover, the main focus of the analysis was to describe the usage patterns of these auxiliaries.

***Sample size and representativeness*** Before starting the analysis, it is important to address a few key details regarding the representativeness of the sample used. A dependable and robust model must consider a broad spectrum of sentences to provide a reliable explanation for the sentences included in the corpus.

In contrast to non-peripheral verbs, peripheral verbs exhibit a wide range of contexts of use. Therefore, the CQL query on Sketch Engine was used to obtain a sample covering the full spectrum of possible contexts, ensuring robustness and reliability in accurately predicting the auxiliary usage of a given verb in a given sentence. A random sample alone may not have guaranteed coverage of all potential occurrences. The CQL query function provided by Sketch Engine addresses this issue by allowing us to target specific verbs for analysis, thus reducing inherent corpus variability.

An illustration of the challenge in capturing the specificities of each verb is provided by a frequency analysis represented by the contingency table. Among the analyzed verbs, the contingency table on the first 100 occurrences extracted from ItTenTen (16) for each verb may indicate the overall frequency of the verbs, but it does not capture the specific contexts. For

instance, there are verbs where the situation is clear and there is a statistically significant differences in auxiliary selection.

Consider the verb *cambiare* as an example:

Contingency Tables			
	TRANS		
AUX	0=intransitive	1=transitive	Total
0=have	1	21	22
1=be	78	0	78
Total	79	21	100

TABLE 1 – Contingency table for *cambiare*

In the first 100 occurrences of the verb *cambiare*, using the CQL [tag= “NOUN|PRO:pers”] [lemma= “avere|essere”] [word= “cambiato|cambiata|cambiati|cambiate”], 78 intransitive constructions selecting ‘be’ and 21 transitive constructions selecting ‘have’ were observed. The calculated  $p$ -value in this case is  $p < 0.001$ . This significant difference demonstrates that ‘be’ is exclusively used for intransitive constructions, while ‘have’ emerges as the predominant auxiliary for transitive constructions.

Therefore, the null hypothesis, which posits no association between the verbal construction (intransitive vs. transitive) and the choice of the auxiliary, is rejected, and the alternative hypothesis (H1) is accepted: there is indeed an association between the construction and the auxiliary, where the intransitive constructions select ‘be’ and the transitive constructions select ‘have’.

Although the prevailing tendency leans towards the selection of ‘be’ as the auxiliary in intransitive constructions, the presence of ‘have’ in one occurrence among the first 100 sentences, as well as in the initial analysis, prompts further investigation into the factors influencing its usage.

Additionally, this analysis becomes more complex when applied to verbs that are compatible with both transitive and intransitive constructions but are exclusively considered in relation to one particular meaning associated with one of the constructions. An example of this is *suonare*, which in this study is solely examined in relation to the meaning associated with its intransitive construction as a verb of sound emission while constructions with the verb meaning ‘play’ are not considered. Therefore, a generic contingency table analysis would not capture these differences.

With most of the verbs, the contingency table analysis did not provide enough evidence of the auxiliary preferred due to the specific context addressed in this study for some of the verbs in question. This can be attributed to the nature of peripheral verbs, particularly in Italian. These verbs may use both auxiliaries in the same syntactic context considered, implying that the same intransitive construction can be associated with both ‘have’ and ‘be’, and the initial difference is therefore not statistically significant.

For this reason, a qualitative corpus analysis is valuable for detecting the influence of factors on auxiliary distribution. Once this analysis has been conducted, a statistical analysis will be even more effective.

### 3.1. CHANGE-OF-STATE: *Cambiare*

The verb *cambiare* ‘change’ belongs to the category of *change-of-state* verbs, specifically denoting processes involving a transformation or alteration in the state or condition of an entity. *Cambiare* is a labile verb, meaning it can be used both transitively and intransitively without changing its meaning (e.g., *Giovanni ha cambiato la musica* ‘Giovanni changed the music’ and *La musica è cambiata* ‘The music changed’).

In its intransitive use, this verb exhibits a higher frequency in combination with the auxiliary *essere* ‘be’ in ItTenTen (16): after applying the filter part-of-speech filter ‘no noun in the first five words right’, the proportion between the auxiliaries ‘be’ and ‘have’ amounts to 95% vs 5%, with 14,204 occurrences of ‘be’ and 802 occurrences of ‘have’.

***Human and non-human subjects*** Upon initial observation, there is a correlation between the nature of the subject and the selection of the auxiliary. Non-human subjects are associated with the auxiliary ‘be’, whereas human subjects tend to correlate with ‘have’, as illustrated in Figure 3.1.

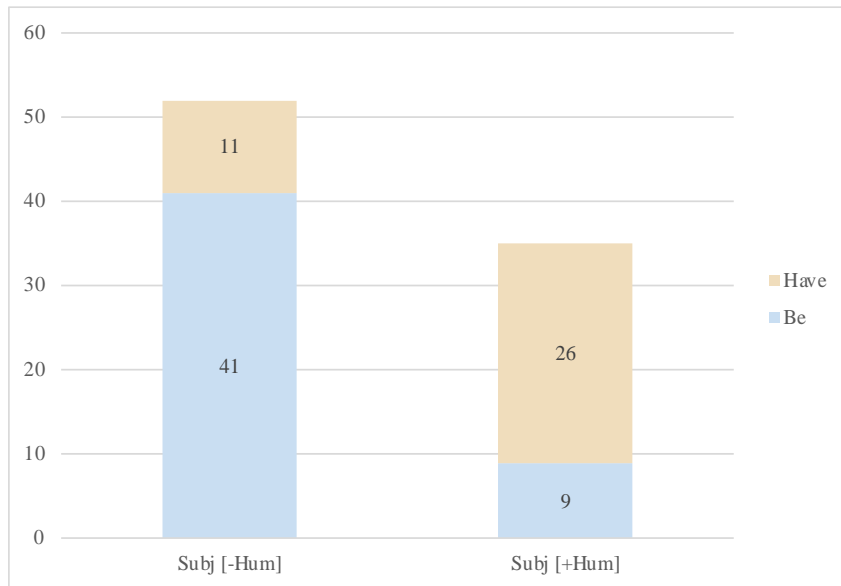


FIGURE 3.1 – Correlation of ‘have’ and ‘be’, and the human/non-human nature of the subject in *cambiare*.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	41	11	52
Subj [+Hum]	9	26	35
<b>Total</b>	50	37	87

TABLE 3.1 – Frequency of ‘have’ and ‘be’ according to the human/non-human nature of the subject.

As depicted in Table 3.1, there are 87 tokens analyzed with *cambiare* instead of the intended 100, as it was not possible to locate a sufficient number of intransitive sentences with ‘have’. When occurrences that potentially imply a direct object complement in the intransitives are distinguished, the occurrences selecting ‘have’ for human subjects when *cambiare* is intransitive decrease (Figure 3.2).

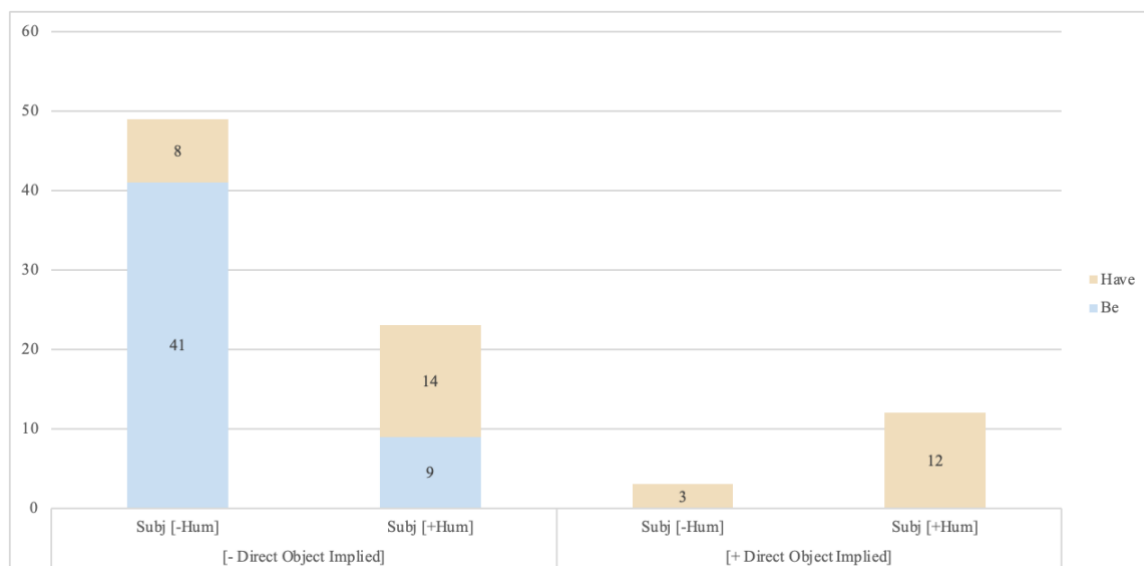


FIGURE 3.2 - Correlation of ‘have’ and ‘be’, the human/non-human nature of the subject in the intransitive use of *cambiare*.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] Subj [-Hum]</b>	41	8	49
<b>[-DO] Subj[+Hum]</b>	9	14	23
<b>[+DO] Subj [-Hum]</b>	/	3	3
<b>[+DO] Subj [+Hum]</b>	/	12	12
<b>Total</b>	50	37	87

TABLE 3.2 - Frequency of ‘have’ and ‘be’ according to the human/non-human nature of the subject in the intransitive use of *cambiare*

As demonstrated in Figure and Table 3.2, when an object is involved, whether implicit or explicit, only ‘have’ is accepted for both non-human (1) and human subjects (2):

- (1) Il tempo ha cambiato tanto, e anche me.  
 DEF.ART time have.PRS.3SG change.PTCP much/many things and even me

‘Time has changed so much/many things, and even me.’

- (2) [...] negli ultimi tempi ho cambiato un po’ [...].  
 in. DEF.ART recent years have. PRS.1SG change.PTCP a.little

‘[...] in recent times I’ve changed a little [...].’

In (1), *tanto* is employed as a quantifying adverbial (‘many things’), similar to *un po’* in (2). Despite potential ambiguity, the subject appears to be more of a willing cause of the change than an entity impacted by it. Consequently, *un po’* ‘a little’ is interpreted as ‘some things’, and the constructions are considered transitive.

When the verb is used intransitively, non-human subjects favor ‘be’ (3) and human subjects ‘have’ (4):

(3) La mia vita è cambiata, la mia postura è  
DEF.ART My life be- PRS.3SG change.PTCP.F.SG DEF.ART my posture be. PRS.3SG

cambiata.  
change.PTCP.F.SG

‘My life has changed, my posture has changed.’

(4) Nel frattempo ho cambiato ancora con un netgear e  
Meanwhile have- PRS.1SG change.PTCP again with INDF.ART netgear and

ora invece uso un asus [...].  
now instead use. PRS.1SG INDF.ART asus

‘[...] meanwhile I changed again to a netgear, and now I use an asus instead [...].’

In (3), the non-human subject represents an entity undergoing a change (‘my life has changed’), whereas in (4), the human subject exhibits agentivity even within an intransitive context. The subject in (4) has chosen to make a change and switch to a neatgear. Hence, upon considering agentivity, it becomes evident that ‘be’ is favored due to the non-agentive nature of the subject. Conversely, with human subjects, ‘have’ is preferred when the subject exhibits agency.

**Agentivity** When considering the factors of agentivity and internal cause, the following pattern emerges: non-agentive subjects select ‘be’, while agentive ones ‘have’, and internal cause renders both auxiliaries possible (Figure 3.3).

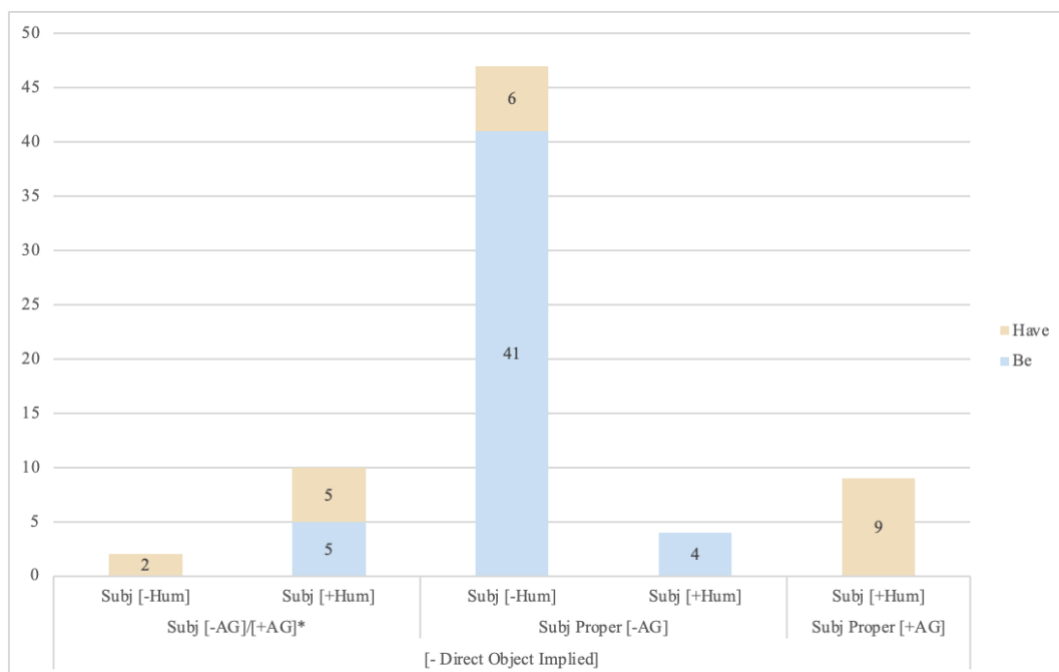


FIGURE 3.3 – Correlation between ‘have’ and ‘be’, agentivity, internal cause<sup>13</sup>, lack of agency, and subject nature (human/non-human) in the intransitive use of *cambiare*.

Factors	Data ‘Be’	Data ‘Have’	Total
[-DO] Subj [-AG] [-AG]/[+AG]*[-Hum]	/	2	2
[-DO] Subj [-AG] [-AG]/[+AG]*[+Hum]	5	5	10
[-DO] Subj [-AG] proper [-AG][-Hum]	41	6	47
[-DO] Subj [-AG] proper [-AG][+Hum]	4	/	4
[-DO] Subj [+AG] proper [+AG][+Hum]	/	9	9
<b>Total</b>	50	22	72

TABLE 3.3 – Frequency of ‘have’ and ‘be’ according to the nature of the subject in terms of agentivity, internal cause, lack of agency, and subject nature (human/non-human).

As clearly indicated in Table and Figure 3.3, the semantic factors of agentivity/lack of agentivity/internal cause play a crucial role in the selection of the auxiliary in the intransitive use of *cambiare*. When the subject lacks agentivity, the auxiliary ‘be’ is employed for both non-human (5) and human subjects (6):

<sup>13</sup> Instances of *Internal causation*= Subj [-AG]/[+AG] \* are distinguished from proper non-agentive subjects.

(5) [...]la mia vita è cambiata da quando ho  
 DEF.ART my life be.PRS.3SG change.PTCP.F.SG since when have.PRS.1SG

deciso di interessarmi a questa vicenda.  
 decide.PTCP to take.an.interest-REFL to this matter

‘[...] my life has changed since I decided to take an interest in this matter.’

(6) Il pubblico d’altra parte è cambiato, è molto più giovane.  
 DEF.ART audience on.other hand be- change- be- molto più giovane  
 PRS.3SG PTCP PRS.3SG much more young

‘The audience on the other hand has changed, is much younger.’

In both (5) and (6), the subjects are wholly affected by the change. In (6), despite being human, the subject undergoes a change, as the audience is not responsible for becoming younger.

When subjects represent an internal cause, it is challenging to ascertain the extent to which they are affected by the change versus being the cause of it, especially due to the low number of occurrences. These instances are classified as potential causative (hence denoted as [-AG]/[+AG]\*). In this context, non-human subjects typically use ‘have’ (7):

(7) woowoow flashgames ha cambiato ...era da un po’  
 wow flashgames have. PRS.3SG change.PTCP be.IPFV.3SG since INDF.ART while

che non venivo più... speriamo che anche i  
 that not come IPFV.1SG more hope.PRS.1PL that even DEF.ART

giochi siano migliorati<sup>14</sup>.  
 games have.SUBJ.3PL improve.PTCP.M.PL

‘wow flashgames has changed...it’s been a while since I’ve been here...hopefully the games have improved as well.’

As previously discussed, (§2.1.3), the subject in example (7) is potentially causative due to interpretative ambiguity: on one hand, the website ‘flashgames’ could be understood as the affected entity upon which an external agent performed the action. On the other hand, it may suggest that the change was executed by others but is attributed to the website, reflecting a form of *shifted intentionality*. Additionally, the use of ‘have’ may introduce an element of circular reasoning, which can be addressed by adopting a construction grammar framework (Goldberg 1995).

<sup>14</sup> Because this is colloquial language taken from forums, some typos have been corrected in the present example.

Returning to our analysis, when the subject consists of internal cause but is human, albeit with limited data, we observe the occurrence of both ‘have’ (8) and ‘be’ (9):

- (8) Ma mai l’animale-uomo ha cambiato, e cambierà,  
 but never DEF.ART.animal-man have.PRS.3SG change.PTCP and change.FUT.3SG

nelle sue forme di sostentamento.  
 in. DEF.ART its forms of sustenance

‘But never has the animal-man changed, and will change, in its forms of sustenance’

- (9) Invece io sono cambiata e lei non lo vuole capire [...].  
 instead I be.PRS.1SG change.PTCP.F.SG and she not it want.PRS.3SG understand.INF

‘Instead, I have changed and she does not want to understand it [...].’

In (8) and (9), the two human subjects could potentially be the cause of the change (even voluntarily), but they could also be entities affected. Consequently, they are not solely agentive but rather potentially causative.

In Italian, the canonical auxiliary employed with *cambiare* would be ‘be’, as evidenced by the predominance of instances selecting ‘be’, except for when the subject presents agentivity. Nevertheless, the concept of internal causality can elucidate the residual acceptance of ‘have’ in certain cases.

Other factors complicating the analysis are the presence of subjects that are somewhat animate and human via metonymy, such as (10):

- (10) Il suo sguardo ha cambiato e da sospettoso è  
 DEF.ART his gaze have.PRS.3SG change.PTCP and from suspicious be.PRS.3SG

diventato curioso.  
 become.PTCP curious

‘His gaze has changed and from suspicious became curious.’

In (10) *sguardo* ‘gaze’ is a metonym for the person looking and it is impossible to identify whether they are rather *affected by change* or are rather *causes of change*. This sentence is particularly significant as it underscores the challenges encountered during the analysis due to subjects that appear to be non-human but are actually human through metonymy (see Flaux & Van de Velde 2000 for a further analysis of hard-to-classify names).

**Adverbials** The most common adverbials are those of manner and time, yet the adverbial of quantity, while less frequent, is also noteworthy for analysis.

**ADVERBIALS OF MANNER** The adverbials of manner, such as *caratterialmente* ‘temperamentally’, total 12 occurrences, evenly distributed with 6 instances each with ‘have’ and ‘be’. This distribution is evident except for occurrences with non-agentive subjects that select ‘be’. The pattern of adverbial of manner with ‘have’ and ‘be’ is observed particularly when the subject is human and internal cause (11; 12; 13):

- (11) Ci   sentivamo   spesso   durante   il           giorno   e   poi   un           giorno  
REFL   talk.IPFV.1PL   often   during   DEF.ART   day   and   then   INDF.ART   day
- ha                   cambiato           completamente.  
have.PRS.3SG   change.PTCP   completely

‘We used to talk a lot during the day, and then one day he completely changed.’

- (12) Con   qualche   dritta   e   l’aiuto       di   esperti   ho           cambiato  
with   some   tip   and   DEF.ART. help   of   experts   have.PRS.1SG   change.PTCP

‘da così a così.’  
from.this.to.that

‘With some tips and the help of experts, I changed *from this to that*.’

- (13) Nel       frattempo   È           cambiato   anche   caratterialmente [...].  
in.DEF.ART   meantime   be.PRS.3SG   change.PTCP   event   character.wise

‘In the meantime, he has also changed character-wise [...].’

In (11) and (12), occurrences with ‘have’ can be elucidated by the causative interpretation of the subject. Furthermore, (11) closely resembles (13), offering evidence for the ambiguous reading facilitated by the causative subject, thereby permitting the option of two auxiliaries.

While the data may be limited, a clear trend has emerged from the examples examined thus far. On the one hand, subjects conveying the feature of internal cause demonstrate a preference for ‘have’. On the other hand, the presence of ‘have’ aids in attributing internal cause to the subject, particularly when the subject is human.

The adverbial of manner combined with ‘have’ can lead to the interpretation of subjects as internal cause, although there are cases such as (14) where the internal cause interpretation is impossible. In latter instance, the subject is non-human and non-agentive.

- (14) [...] sottolineando soprattutto gli aspetti di lui/lei che amavi e  
 emphasizing especially DEF.ART aspects of him/her that love.IPFV.2SG and
- che si sono persi con l'assunzione della droga, e  
 that REFL be.PRS.3PL loose-PTCP.M.PL with DEF.ART.taking of.DEF.ART drug and
- come la sua vita abbia cambiato solo in peggio.  
 how DEF.ART his/her life have.SUBJ.3SG change.PTCP only in worse
- '[...] especially emphasizing the aspects of him/her that you loved that were lost with taking drugs, and how his/her life changed only for the worse.'

- (15) La mia soddisfazione più grande è rivedere un  
 DEF.ART my satisfaction more great be- see.again- INDF.ART  
 PRS.3SG INF
- cliente e sentirmi dire quanto la sua vita  
 client and hear.me.INF say.INF how.much DEF.ART his life
- sia cambiata in meglio.  
 be.SUBJ.3SG change.PTCP.F.SG in better
- 'My greatest satisfaction is seeing a client again and hearing how much his life has changed for the better.'

The canonical auxiliary in this scenario, involving a non-human non-agentive subject, is 'be', as exemplified by (15).

**QUANTITY ADVERBIAL** In the presence of an adverbial of quantity, such as *molto* 'much', the sole auxiliary found in this corpus is 'have'. The preference for 'have' could be attributed to the fact that this quantity adverbial can occur in two constructions: a transitive one (16) and an intransitive one (17;18):

- (16=1) Il tempo ha cambiato tanto, e anche me.  
 DEF.ART time have. PRS.3SG change.PTCP much/many things and even me
- 'Time has changed so much/many things, and even me.'

- (17) La storia ha cambiato un po', invece di salvare le  
 DEF.ART story have. PRS.3SG change.PTCP a.bit instead of save.INF DEF.ART
- uova di Yoshi, Mario dovrà trovarli.  
 eggs of Y. M. need.FUT.3SG find.them.INF
- 'The story has changed a bit, instead of saving Yoshi's eggs, Mario will have to find them.'

- (18) Durante il Cretaceo, terzo e ultimo periodo, la situazione  
 during DEF.ART Cretaceous third and final period DEF.ART situation  
 ha cambiato molto, sia pure gradatamente.  
 have.PRS.3SG change.PTCP a.lot even.if gradually

‘During the Cretaceous, third and final period, the situation changed greatly, even if gradually’

(16) displays a transitive sentence (and therefore an internal cause subject), as the subject ‘time’ has effectively changed many things, including a person (‘me’). In (18), at first, *molto* can be mistaken for a probable object complement, which leads to an analysis of the subject as causative. However, the adverb that follows, *gradatamente* ‘gradually’, specifies the ongoing change of the entity and states that it is an impacted entity, implying that the subject is not properly agentive, and that the sentence is intransitive. The same analysis applies to (17).

To summarize, even if the canonical auxiliary for intransitive *cambiare* ‘change’ is ‘be’, even with the quantity adverbial (outside the corpus), when combined to a quantity adverbial ‘have’ can appear too. Moreover, the presence of the adverbial of quantity that can often be interpreted as object (e.g., 16=1) and yield a transitive construction for *cambiare*, selecting invariably ‘have’.

**TIME ADVERBIALS** When it comes to temporal adverbials, as *da allora* ‘since then’, it appears that they do not play a role in influencing the auxiliary distribution of *cambiare* which predominantly selects ‘be’. The sentences featuring ‘be’ typically have non-agentive, non-human subjects.

Therefore, it appears that temporal location does not determine the choice of auxiliary in these cases. Consider (19) as an example:

- (19) [...] oggi la situazione è cambiata [...].  
 today DEF.ART situation be.PRS.3SG change.PTCP.F.SG

‘Today, the situation has changed [...].’

In terms of aspectual adverbials, (20) contains an aspectual adverbial indicating progression, whereas the aspectual adverb in (21) indicates iteration.

- (20) Poi col tempo le cose sono cambiate [...].  
 Then with.DEF.ART time DEF.ART things be.PRS.3PL change.PTCP.F.PL

‘Then over time, things changed.’

- (21) Con la nuova gestione fa schifo! Poi non so quante  
 with DEF.ART new management suck-PRS.3SG then not know-PRS.1SG how.many  
 volte ha cambiato, ma [...] ci sono andato poco.  
 times have-PRS.3SG change.PTCP but LOC be.PRS.1SG go-PTCP little

‘With the new management it sucks! Then I don’t know how many times it changed, but [...]I went there very little [...].’

However, auxiliary selection can be accounted for by the features conveyed by the subject: in (20) the subject is non-agentive and therefore ‘be’ is selected. In (21), the subject (*gestione*) corresponds to an internal cause, and hence the presence of ‘have’ is expected.<sup>15</sup> The nature of the aspectual adverbial does not alter this selection.

To conclude, the most influential factors in the selection of ‘have’ and ‘be’ for a *change-of-state* verb as *cambiare* is the agentivity factor which correlates with the human trait. When the subject is agentive and human, the auxiliary chosen is ‘have’, when the subject is non-agentive, whether human or non-human, the auxiliary is ‘be’. Most of the variation occurs when the subject is internal cause which allows ‘be’ and ‘have’.

Table 3.4 summarizes the principal factors influencing the selection of the auxiliary:

Agentivity	Human Trait	Auxiliary
Agentive	Human →	Have
Internal Cause	Human/Non-human →	Have/Be
Non-Agentive	Human/Non-human →	Be

TABLE 3.4 - Primary factors co-occurring in the influence of the auxiliary for *cambiare*

The two semantic factors, agentivity and human trait, are significant after excluding the transitive construction. Agentivity and human trait mutually influence each other: an internal

<sup>15</sup> The full sentence is: *Rolling Stone... ci sono andato due o tre volte. Una volta a sentire Pino Scotto, sotto Pasqua, con parecchi ospiti più Big di lui, costretto a cedere il passo a un dj set di rock commerciale. Probabilmente io sono della generazione in cui si dicevano cose del tipo "Con la nuova gestione fa schifo!" più o meno ogni anno. Poi non so quante volte ha cambiato, ma come ti dicevo, ci sono andato poco. Il tuo nome l'ho trovato in due secondi, ma sapevo dove cercare!* (Rolling Stone--I went there two or three times. Once to hear Pino Scotto, under Easter, with several guests more ‘Big’ than him, forced to give way to a commercial rock DJ set. I’m probably of the generation where people used to say things like "With the new management it sucks!" pretty much every year. Then I don’t know how many times it changed, but as I told you, I went there very little. Your name, I found it in two seconds, but I knew where to look!). Despite the mention of a bar, we can infer from the sentence that the subject is *gestione* ‘management’.

cause subject tends to favor ‘have’, and the human trait, combined with ‘have’, facilitates the internal cause interpretation.

### 3.2 CHANGE OF POSSESSION: *Cedere*

*Cedere* ‘yield’ ‘surrender’ is a *change-of-possession* verb, denoting the transfer of ownership or control of something from one entity to another. This transfer may involve physical or intangible assets and can encompass acts of yielding or surrendering.

Analyzing the data retrieved from the CQL search on ItTenTen (16), it becomes evident that *cedere* is primarily associated with the auxiliary verb ‘have’. Specifically, after applying the filter part-of-speech filter ‘no noun in the first five words right’, the proportion between the auxiliaries ‘have’ and ‘be’ amounts to 88% vs 12%, with 799 occurrences of ‘have’ and 109 of ‘be’.

This preference for the auxiliary ‘have’ is further supported by the subsequent analysis, which could only be conducted on 62 sentences that respected the intransitive requirements. Among these, 50 sentences featured ‘have’ as the auxiliary, while ‘be’ was observed in only 12 instances. However, it is interesting to compare the occurrences of ‘have’ and ‘be’, starting from occurrences with human and non-human subjects for a more comprehensive understanding.

***Human and non-human subjects*** The predominant auxiliary observed is ‘have’, regardless of whether the subject is human or non-human. However, non-human subjects exhibit a greater propensity to accept ‘be’ compared to human subjects. This observation is supported by the data in Figure 3.4 and Table 3.5.

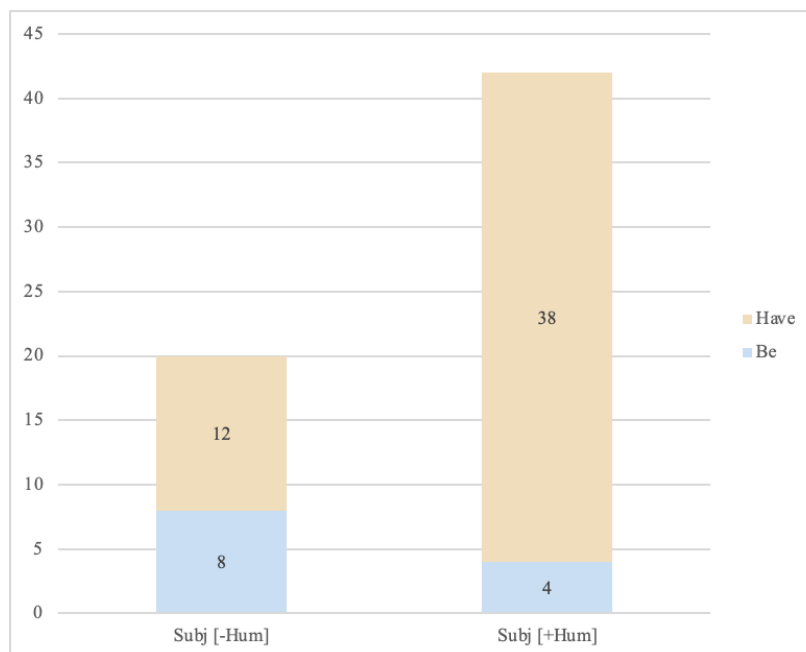


FIGURE 3.4 – Correlation of ‘have’ and ‘be’ and the subject’s human/non-human nature.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	8	12	21
Subj [+Hum]	4	38	41
<b>Total</b>	12	50	62

TABLE 3.5 – Frequency of ‘have’ and ‘be’ according to the subject’s human/non-human nature.

As shown in Figure 3.4. and Table 3.5, the majority of the data consist of occurrences with human subjects, totaling 41 instances. In both human and non-human subjects, ‘have’ predominates (22; 23). However, some non-human subjects exhibit occurrences of ‘be’ (24):

- (22) Ma me ne ero innamorato e alla fine ho  
 But REFL of.her be.IPFV.1SG fall.in.love.PTCP and finally have. PRS.1SG

ceduto.  
 yield. PTCP

‘But I fell in love with it and finally yielded.’

- (23) Ormai l’argine ha ceduto.  
 by.now DEF.ART.levee have.PRS.3SG yield.PTCP

‘By now the levee has yielded.’

- (24) Una lastra È ceduta sotto il suo peso.  
 INDF.ART slab be.PRS.3SG yield.PTCP.F.SG under DEF.ART his weight

‘A slab is yielded under his weight.’

In sentence (22), the meaning of *cedere* is interpreted as ‘give in’ or ‘give in to temptation’, a reading that recurs with human subjects. Conversely, sentences (23) and (24) demonstrate the meaning most commonly associated with non-human subjects, namely ‘yield’ or ‘collapse’.

Furthermore, these sentences reflect a distinction in agentivity: the interpretation in (22) is applicable to both agentive and internal cause subjects, while the readings in (23) and (24) necessitate non-agentive subjects.

**Agentivity** Agentivity holds considerable significance: when the subject is either agentive or internal cause, the auxiliary verb selected is typically ‘have’. In cases where the subject is non-agentive, ‘have’ remains prevalent, but ‘be’ becomes a possible alternative, as illustrated in Figure 3.5.

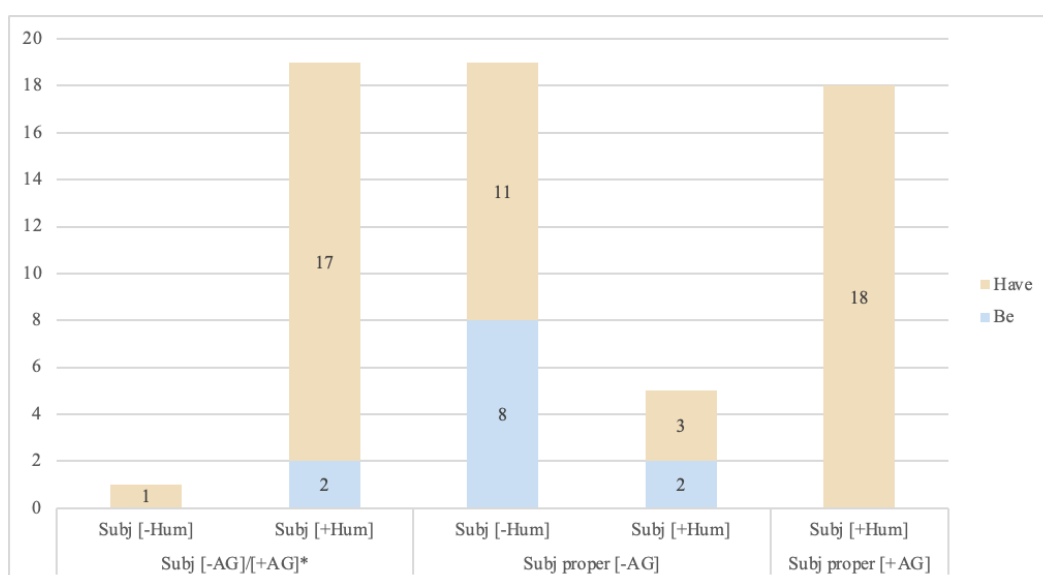


FIGURE 3.5 – Correlation of ‘have’ and ‘be’, agentivity and the subject’s human/non-human nature.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG] [-AG]/[+AG] *[-Hum]	/	1	1
Subj [-AG] [-AG]/[+AG] *[+Hum]	2	17	19
Subj [-AG] proper [-AG] [-Hum]	8	11	19
Subj [-AG] proper [-AG] [+Hum]	2	3	5
Subj [+AG] proper [+AG] [+Hum]	/	18	18
<b>Total</b>	12	50	62

TABLE 3.6 – Frequency of ‘have’ and ‘be’ according to agentivity and the subject’s human/non-human nature.

When the subject is agentive (and, therefore, human), the interpretation of *cedere* is that of voluntarily ‘give in’, and the auxiliary selected is ‘have’, as exemplified in sentence (25).

- (25) Ce l'avevo in testa da un po'. Alla fine ho ceduto  
 this it.have- IPFV.1SG on mind since a.while finally have.PRS.1SG give.in.PTCP  
 e me lo sono preso.  
 and REFL it be.PRS.1SG get.PTCP

'It had been on my mind for a while. I finally gave in and got it.'

In sentence (25), the action is evidently voluntary, as highlighted by the subsequent sentence *me lo sono preso* 'I got it', which signifies a certain degree of agency.

However, differently to what emerges in Sorace (2000: 874-875) about *cedere*, the verb *cedere* does not always 'imply deliberate intentionality on the part of the subject'. On the contrary, it appears that the usage of *cedere* could imply intrinsically a certain lack of agentivity that converts into agentivity in specific scenarios, as demonstrated in (25). In general, when one yields to something, it suggests a lack of volition, indicating that the action was undertaken without a desired intention, but rather due to being left with no alternative but to concede. At the same time, the act of yielding is a choice.

There are also occurrences where the subject serves as an internal cause, indicating both control and patient-like characteristics simultaneously. In this scenario, 'have' prevails (26):

- (26) Quegli stupidi dei miei occhi hanno ceduto, è  
 those stupid of.DEF.ART my eyes have. PRS.3PL give.in.PTCP be.PRS.3SG  
 durato solo un attimo [...].  
 last.PTCP only INDF.ART moment

'Those stupid eyes of mine gave in; it lasted only a moment [...].'

In sentence (26), *occhi* 'eyes' is considered human since it represents a human entity through metonymy. This sentence illustrates how the subject's eyes are causative without necessarily implying volition: they resist yielding but eventually succumb to the temptation to stare.

The distinction is clear in occurrence (27), where a human subject is presented through metonymy but is not inherently agentive.

- (27) Poi purtroppo il suo robustissimo fisico ha ceduto.  
 then unfortunately DEF.ART his very.robust body have.PRS.3SG collapse.PTCP  
 È mancato alle 5 di mattina [...].  
 be.PRS.3SG pass.PTCP at. DEF.ART 5 of morning

'Then unfortunately his very robust body also gave out. He was passed away at 5 a.m. [...].'

In sentence (27), *fisico* is definitely non-agentive because the subject has passed away (*è mancato alle 6 di mattina*), and there is no internal control or inherent property determining the action. This pattern is consistent across all occurrences of human non-agentive subjects represented by body parts, as exemplified in (27). The interpretation associated with constructions featuring a non-agentive subject is that of ‘collapse’, as evident from non-human subjects, which can select both ‘have’ (28;30) and ‘be’ (29;31).

In other words, *cedere* is predominantly a ‘have’ selection verb, as indicated by the analysis, but the absence of agentivity allows for the appearance of ‘be’.

(28) La        lastra    di    marmo    del        balcone    ha        ceduto.  
 DEF.ART   slab     of    marble    of. DEF.ART   balcony    have.PRS.3SG    yield.PTCP

‘The marble slab on the balcony yielded.’

(29=24) Una        lastra    è        ceduta        sotto    il        suo    peso.  
 INDF.ART   slab     be-.PRS.3SG    yield.PTCP.F.SG    under    DEF.ART   his    weight

‘A slab is yielded under his weight.’

(30) Sotto    una        delle        sue    ruote    anteriori    l’asfalto        ha  
 under    INDF.ART   of. DEF.ART   its    wheels    front        DEF.ART. asphalt    have.PRS.3SG  
 ceduto.  
 yield.PTCP

‘Under one of his front wheels the asphalt collapsed.’

(31) [...]in    alcuni    punti [...]    il        sentiero    è        ceduto.  
 in        some    places        DEF.ART    path        be.PRS.3SG    yield.PTCP

‘[...] in some places the path has yielded [...].’

The corpus analysis proves to be essential as it has uncovered a significant aspect: the verb *cedere* predominantly selects ‘have’ as its auxiliary due to the presence of agentivity or internal cause in the subject. Consequently, the findings presented in Sorace (2000: 874-875) are only partly consistent. Sorace notes that “some of these verbs can appear with a non-human subject: when they do, they also tend to be acceptable with the ‘be’ auxiliary.” While ‘be’ is sometimes acceptable, *cedere* indeed more readily associates with ‘have’. However, this tendency is particularly observed with agentive and internal cause subjects, which are typically human or closely linked to a human referent, presenting a specific reading such as ‘yield’ or ‘give in’.

When the subject lacks agentivity, both auxiliaries are possible. On the one hand, ‘have’, typically associated with the verb *cedere* may be used. On the other hand, *cedere* meaning ‘collapse’ when the subject is non-human

**Adverbials** There are not many occurrences featuring adverbials with *cedere*. The most frequent one is the time adverbial.

**TIME ADVERBIALS** There are 13 occurrences featuring time adverbials such as a *tarda sera* ‘late at night’. Among these 13 occurrences, 11 sentences select ‘have’ and feature human subjects, while the other two occurrences involve non-human subjects and present both ‘have’ and ‘be.’

The predominant selection of ‘have’ and the possibility of both ‘have’ and ‘be’ (although the occurrences are insufficient to make a precise estimate) with non-human subjects align with the general tendency observed with *cedere*. As a result, it seems that the temporal adverbial has no significant effect on the choice of auxiliary.

To conclude, the verb *cedere* is significantly influenced by the semantic factor of agentivity. When embedded in a sentence with an agentive or internal cause subject, particularly if human, the preferred auxiliary is ‘have’. This analysis underscores the tendency of *cedere* to select ‘have’.

Simultaneously, when the subject lacks agentivity, both ‘have’ and ‘be’ are possible options, particularly with non-human subjects.

Agentivity	Human Trait	Auxiliary
Agentive, internal Cause	Human →	Have
Non-Agentive	Human/Non-human →	Have/Be

TABLE 3.7 – Primary factors affecting the selection of the auxiliary in *cedere*.

### 3.3 VERBS OF SOUND EMISSION: *suonare*

*Suonare* ‘ring’ belongs to the class of verbs of sound emission, which encompasses verbs describing the production or emission of sounds. Analysis conducted on the ItTenTen (16) corpus using the Corpus Query Language (CQL) reveals a comparable number of occurrences of ‘have’ and ‘be’ in this context. Specifically, applying the filter to restrict occurrences where a noun does not appear within the first five tokens, the auxiliary ‘have’ appears 59% of the

time (367 occurrences of ‘have’), while the auxiliary ‘be’ appears 41% of the time (259 occurrences of ‘be’).

**Non-human subjects** The verb *suonare* ‘ring’ is exclusively associated with non-human subjects. It serves as an example of verbs with auxiliary alternation, as it can co-occur with both ‘have’ and ‘be’ in contexts involving sound emission, as evidenced by examples (32) and (33):

- (32) La campana ha suonato è ora di fare  
DEF.ART bell have.PRS.3SG ring.PTCP be.PRS.3SG time to make.INF  
 uno sprint per non finire ultimi.  
INDF.ART sprint to not finish.INF last

‘The bell has rung; it’s time to sprint so we don’t finish last.’

- (33) Poi la campanella è suonata e siamo  
Then DEF.ART bell be.PRS.3SG ring.PTCP.F.SG and be.PRS.1PL  
 dovuti ritornare in classe.  
have.to- PTCP.M.PL return- INF in class

‘Then the bell rang and we had to return to class.’

As demonstrated in examples (32) and (33), both ‘have’ and ‘be’ are employed in analogous contexts with the same subject (the ‘bell’). At first glance, no discernible difference appears, but the allowance for both ‘have’ and ‘be’ stems from the internal cause’s nature of the subject.

The utilization of ‘have’ highlights the causative role attributed to the subject, as it is the entity instigating the emission of sound. Simultaneously, the subject ‘bell’ serves as the conduit through which the sound passes, permitting the use of ‘be.’

**Internal cause and (lack of) agentivity** As noted, the factor of internal cause can elucidate the usage of both ‘have’ and ‘be’, owing to the subject’s intermediate nature between exhibiting causative behavior and being an entity subjected to the action. In such instances, ‘have’ is typically preferred, although ‘be’ remains possible.

Conversely, when the subject is not agentive whatsoever, ‘be’ emerges as the auxiliary of choice. Figure 3.6 illustrates this scenario.

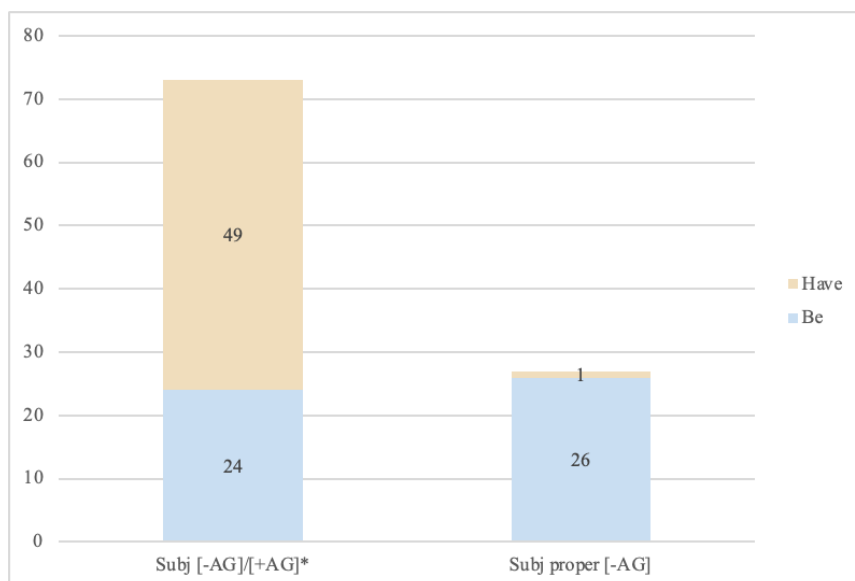


FIGURE 3.6 – Correlation of ‘have’ and ‘be’ and internal cause/lack of agency.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG]/[+AG]*	24	49	73
Subj proper [-AG]	26	1	27
<b>Total</b>	50	50	100

TABLE 3.8 – Frequency of ‘have’ and ‘be’ according to internal cause/lack of agency.

The result aligns with the observations made in examples (32) and (33) and is further illustrated in (34) and (35).

- (34) L’allarme ha suonato alle 2.40 di stanotte.  
 DEF.ART-alarm have.PRS.3SG ring.PTCP at.DEF.ART 2.40 of tonight

‘The alarm rang at 2.40 a.m. tonight.’

- (35) L’allarme del negozio è suonato costringendoli a  
 DEF.ART-alarm of.DEF.ART store be.PRS.3SG ring.PTCP forcing.them to

scappare.  
 leave-INF

‘The store alarm rang forcing them to leave.’

Examples (34) and (35) depict the same subject, ‘alarm’, which acts as an internal cause and exhibits causative behavior without volition. The choice between ‘have’ and ‘be’ depends on the perceived degree of causativity associated with the non-human entity and the impossibility of detecting if it is an agent or a patient.

On one hand, the causativity of the alarm arises from its role as an instrumental entity designed to emit sound. On the other hand, the sentence could portray an event where the emission of sound is attributed to an external cause, mediated through the entity representing the subject, in this case, the ‘alarm’.

In contrast to these instances are those featuring non-agentive subjects, exemplified by (36) and (37). In such cases, only ‘be’ is employed:

(36) Lo sfogo del tecnico salentino è suonato  
 DEF.ART outburst of.DEF.ART technician Salento. ADJ be.PRS.3SG sound.PTCP

familiare.  
 familiar

‘The Salento technician’s outburst sounded familiar.’

(37) Ma io vi dico che il momento è giunto,  
 But I to.you say.PRS.1SG that DEF.ART time be.PRS.3SG come.PTCP

l’ora è suonata.  
 DEF.ART.hour be.PRS.3SG sound.PTCP.F.SG

‘But I say to you that the time has come, the hour has sounded.’

In (36), *suonare* + adjective is a copulative construction close to ‘appear’ + adjective but involving auditive rather than visual perception. (37) features the fixed expression *l’ora è suonata*, which metaphorically signifies the arrival of a moment, as supported by the preceding expression *il momento è giunto* ‘the moment has come’. In both examples, the subject is non-agentive, and the fixed expression can impact the choice of the auxiliary.

**Adverbials** Among the most recurrent adverbials, temporal adverbials and atelic aspectual ones stand out. However, the aspectual adverbials are particularly intriguing in terms of their impact on the selection of ‘have’.

**TIME ADVERBIALS** Time adverbials such as *alle 2.40* ‘at 2.40 a.m.’ are notably common with *suonare* (Figure 3.7).

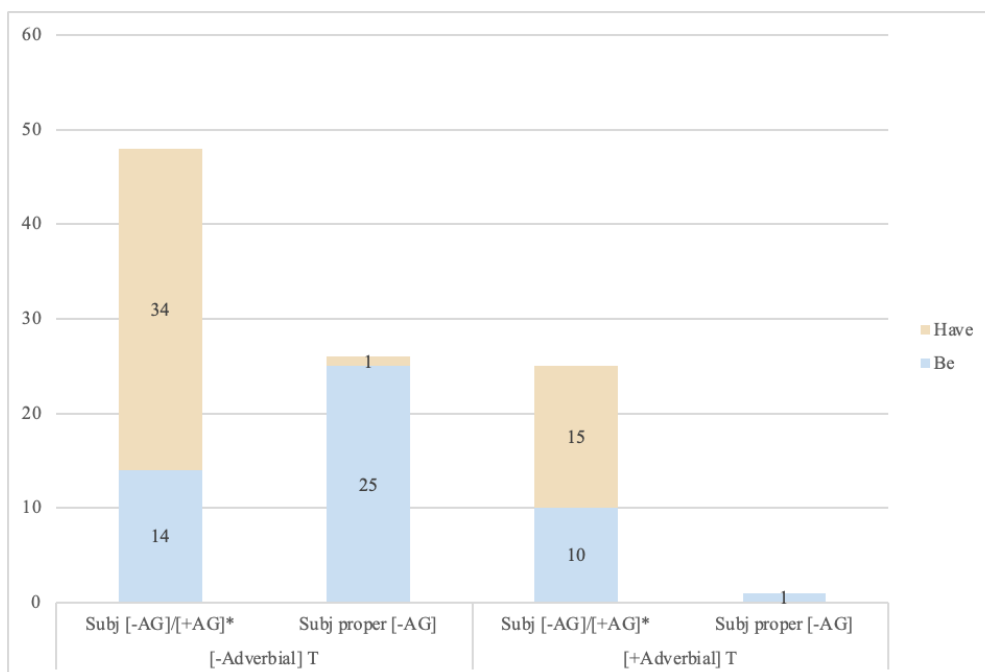


FIGURE 3.7 – Correlation of ‘have’ and ‘be’ with time adverbials and internal cause/lack of agentivity.

Factors	Data ‘Be’	Data ‘Have’	Total
[-Adverbial T] Subj [-AG]/[+AG] *	14	34	48
[-Adverbial T] Subj proper [-AG]	25	1	26
[+Adverbial T] Subj [-AG]/[+AG] *	10	15	25
[+Adverbial T] Subj proper [-AG]	1	/	1
<b>Total</b>	50	50	100

TABLE 3.9 – Frequency of ‘have’ and ‘be’ according to time adverbials and internal cause/lack of agentivity.

Among the 24 occurrences, there is a discernible preference for ‘have’ when the subject functions as an internal cause, as illustrated in Figure 3.7. Figure 3.7 and Table 3.9 demonstrate that the time adverbial does not influence the choice of the auxiliary. The preference for ‘have’ remains higher when the subject serves as an internal cause, regardless of the presence or absence of the time adverbial, as depicted in (38):

- (38) Alle 15:40 il campanello ha suonato.  
 at.DEF.ART 3.40 p.m. DEF.ART doorbell have.PRS.3SG ring.PTCP

‘At 3:40 p.m. the doorbell rang.’

**ASPECTUAL ADVERBIALS** Occurrences featuring aspectual adverbials, particularly those with atelic or progressive meanings, tend to select ‘have’. Both in (39), containing an atelic adverbial, and in (40), with an iterative adverb, the auxiliary ‘have’ is used:

- (39) Il campanello ha suonato a lungo.  
 DEF.ART doorbell have. PRS.3SG ring.PTCP for long

‘The doorbell has rung for long.’

- (40) Oggi l’avviso della mail ha suonato spessissimo.  
 today DEF.ART.alert of. DEF.ART mail have.PRS.3SG ring. PTCP very.often

‘Today the mail alert has rung very often.’

In (39) and (40), both types of aspectual adverbials emphasize the ongoing nature of the verb’s process. Additionally, the sentences feature an internal cause subject, which aligns with the interpretation of an ongoing process.

Even in the absence of aspectual adverbials, an ongoing process can be emphasized by several factors, including the choice of auxiliary itself. For instance, the use of ‘have’ and the repetition of the verb *suonare* can underscore the continuous nature of the process, whereas the use of ‘be’ may suggest the interpretation of an event brought to its completion.

In the following examples, (41) and (42), both feature an internal cause subject.

- (41) [...] quel cellulare ha suonato e suonato e nessuno  
 that cell.phone have-PRS.3SG ring- PTCP and ring- PTCP and nobody

rispose.  
 answer. IPFV.3SG

‘[...] that cell phone rang and rang, and no one answered.’

- (42) Proprio a questo punto è suonata la sveglia.  
 just at this point be.PRS.3SG ring. PTCP.F.SG DEF.ART alarm

‘Just at this point the alarm went off.’

In (41), the emphasis on the action and the process is achieved through the repetition of the verb. Conversely, in (42), the use of ‘be’ and the expression *proprio a questo punto* ‘right at this moment’ steer the interpretation towards viewing the sentence as an event that occurred rather than an ongoing process in relation to the causative action of the subject.

Consequently, along with aspectual adverbials and features such as the repetition of the verb, ‘have’ aligns with the interpretation of a process, while ‘be’ aligns with the portrayal of a completed event.

In conclusion, this verb exhibits a balanced distribution of ‘have’ and ‘be’. When the subject is an internal cause ‘have’ predominates. Conversely, when the subject lacks agency, the verb takes on a metaphorical meaning (e.g., ‘*suonare* + adjective’) and the auxiliary ‘be’ is selected.

Additionally, in the former case, aspect plays a significant role: an ongoing process typically aligns with ‘have’, whereas a resultative one aligns with ‘be’. Table 3.10 summarizes the most important factors:

Agentivity	Aspect	Auxiliary
Internal Cause	Process →	Have
	Result →	Be
Non-Agentive	/ →	Be

TABLE 3.10 – Primary factors affecting the selection of the auxiliary in *suonare*.

### 3.4 EMOTIONAL STATE VERBS

Emotional state verbs are verbs that describe the feelings or emotions of individuals. In the corpus used for this study, among the emotional state verbs, there are *contare* and *pesare*, used in their derived meaning of ‘matter’.

#### 3.4.1 *Contare*

The verb *contare*, which conveys meanings of ‘matter’ or ‘count’, is notably present in the ItTenTen (16) corpus, particularly in conjunction with ‘have’. In the absence of a noun in the first five words following the past participle, it appears with both ‘have’ and ‘be’ in 50% of instances totaling 215 occurrences. However, it is important to consider that this verb has many significations with ‘be’, as in *le risorse sono contate* (‘resources are limited’).

***Human and non-human subjects*** In this corpus, only 9 occurrences were found with ‘be’. Consequently, the auxiliary ‘have’ emerges as the predominant one, with ‘be’ being almost exclusively associated with non-human subjects, as illustrated by Figure 3.8 and Table 3.11.

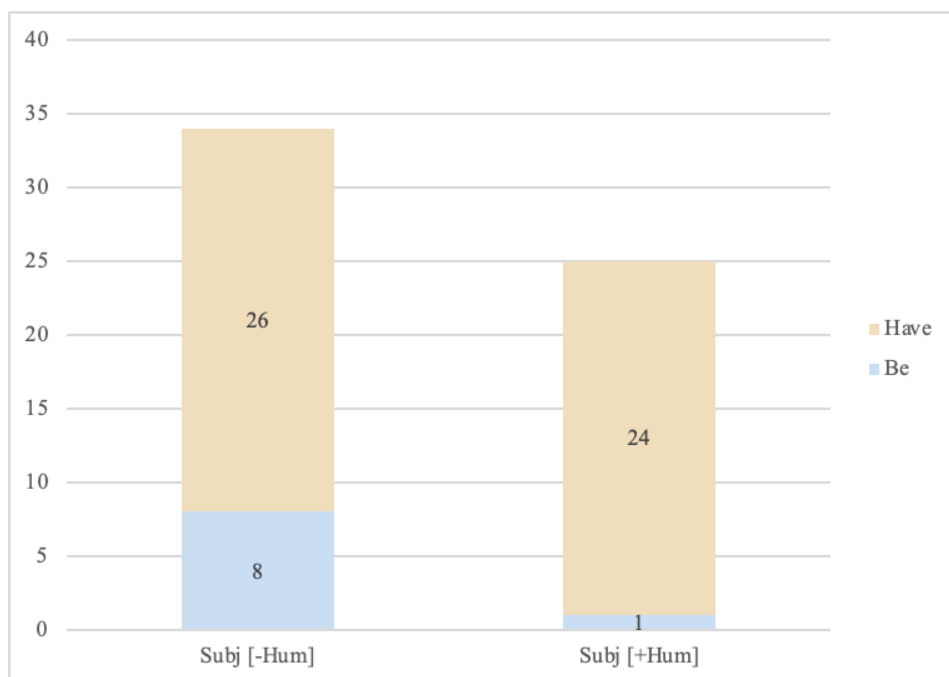


FIGURE 3.8 – Correlation of ‘have’ and ‘be’ and the human/non-human nature of the subject.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	8	26	34
Subj [+Hum]	1	24	25
<b>Total</b>	9	50	59 <sup>16</sup>

TABLE 3.11 – Frequency of ‘have’ and ‘be’ according to the human/non-human nature of the subject.

As depicted in both the figure and the table, human subjects exhibit the auxiliary ‘have’ (with one exception), as exemplified by (43):

- (43) [...] una donna che per lui aveva contato molto.  
INDF.ART woman that for him have. IPFV.3SG matter.PTCP a.lot

‘[...] a [woman] who had meant a lot to him.’

Non-human subjects also display a preference for ‘have’, as demonstrated in (44), but they also admit ‘be’, as illustrated in (45):

- (44) [...] nella sua educazione la guerra ha contato molto.  
in.DEF.ART her education DEF.ART war have.PRS.3SG matter.PTCP a.lot

‘[...] war counted a lot in her education.’

<sup>16</sup> There are 59 occurrences since only 9 of the 215 occurrences found through the CQL with ‘be’ are intransitive.

(45) Tutta l'esperienza ed i contatti [...] sono contati  
 all DEF.ART.experience and DEF.ART contacts be.PRS.3PL matter. PTCP.M.PL  
 e contano tuttora moltissimo.  
 and matter.PRS.3PL still a.lot

'All of your experience and contacts [...] have been and continue to be extremely valuable.'

Both 'have' and 'be' are indeed possible when the subject is non-human, but (45) may sound slightly strange. This becomes clearer when considering agentivity and internal cause: the subject of *contare* is internal cause due to the semantics of the verb. If something or someone matters, they can exert power or control over someone without intending to do so.

Consequently, the subject is inherently causative, whether it acts intentionally as a human subject, or serves as an internal cause. As a result, *contare* is predominantly a verb that selects 'have' due to this internal cause factor.

**Agentivity and internal cause** *Contare* primarily involves internal cause subjects, although some humans can be agentive as well. However, the preference for 'have' persists in all scenarios, as demonstrated by Figure 3.9 and Table 3.12.

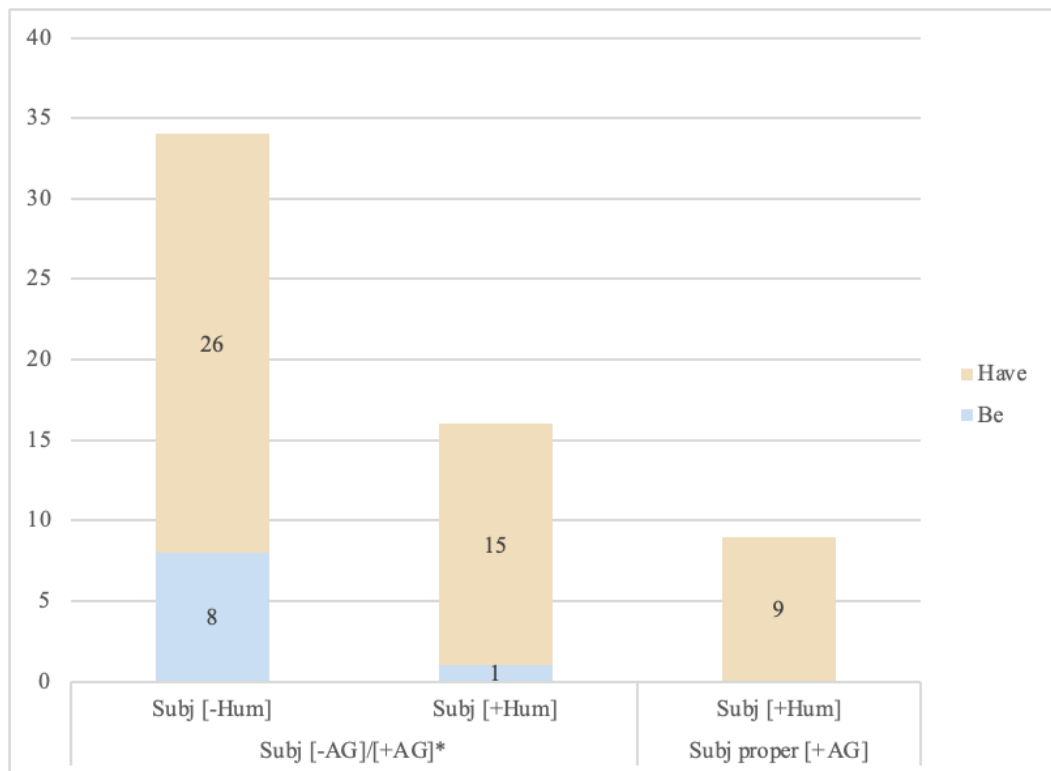


FIGURE 3.9 – Correlation of 'have' and 'be' and the nature of the subject (human/non-human and agentive/internal cause).

Factors	Data 'Be'	Data 'Have'	Total
Subj [-AG] [-AG]/[+AG]*[-Hum]	8	26	34
Subj [-AG] [-AG]/[+AG]*[+Hum]	1	15	16
Subj [+AG] <b>proper</b> [+AG][+Hum]	/	9	9
<b>Total</b>	9	50	59

TABLE 3.12 – Numbers of ‘have’ and ‘be’ depending on the human/non-human and agentive/internal cause nature of the subject.

In our corpus, human subjects can be classified as either non-agentive but potentially causative (46) or properly agentive (47). In all scenarios, ‘have’ is the auxiliary selected:

- (46) Una donna che nella sua vita aveva contato molto [...].  
INDF.ART woman that in. DEF.ART his life have.IPFV.3SG matter. PTCP a.lot

‘A woman who had counted a lot in his life [...].’

- (47) I nostri clienti hanno contato ripetutamente su di noi.  
DEF.ART our clients have.PRS.3PL count.PTCP repeatedly on of us

‘Our customers have repeatedly counted on us.’

The subject (*una donna* ‘a woman’) in (46) is non-agentive but causative as it exerts influence but lacks intentionality. In (47), on the other hand, *clients* repeatedly exercise this voluntary act of trust presenting an indirect transitive use. The comparison between the two is important to understand the difference between a subject characterized by internal cause and an agentive one with *contare*. This comparison highlights that, due to the semantics of the verb, the subject of the intransitive use of *contare* is typically an internal cause.

The selection of ‘be’ is possible with a non-human subject that is not fully agentive, but causative. However, even in such context ‘have’ is preferred (48):

- (48) Ci avete detto che le nostre telefonate avrebbero  
to.us have.PRS.2PL tell.PTCP that DEF.ART our phone.calls have.COND.2PL

contato nel decidere chi avrebbe vinto.  
count.PTCP in.DEF.ART decide.INF who have.COND.3SG win.PTCP

‘You told us that our phone calls would count in deciding who won.’

As demonstrated by (48), *contare* denotes concepts such as ‘have importance’ or ‘have weight’, thereby attributing causative behavior to the subject due to this inherent force. By selecting ‘have’ as an auxiliary, rather than ‘be’ the speaker intends to emphasize the force

exerted by the subject in the act of ‘counting’. Hence, *contare* predominantly selects ‘have’, with exceptions allowing ‘be’, particularly in contexts involving inanimate subjects.

**Adverbials** In the realm of adverbials, both adverbial arguments and quantity adverbials frequently co-occur with *contare*. While these adverbials do not directly dictate the selection of the auxiliary ‘have’ (which remains preferred due to the internal cause), they contribute to underscore the causative essence inherent in the subjects.

**ADVERBIAL ARGUMENTS** *Contare* cooccurs with some complements in a such systematic way that have been considered as arguments of the verb. Among these adverbial arguments, beneficiary (*contare per me* ‘matter to me’) and locative complements (*contare su di me* ‘count on me’) have been distinguished, as Figure 3.10 and Table 3.13 show.

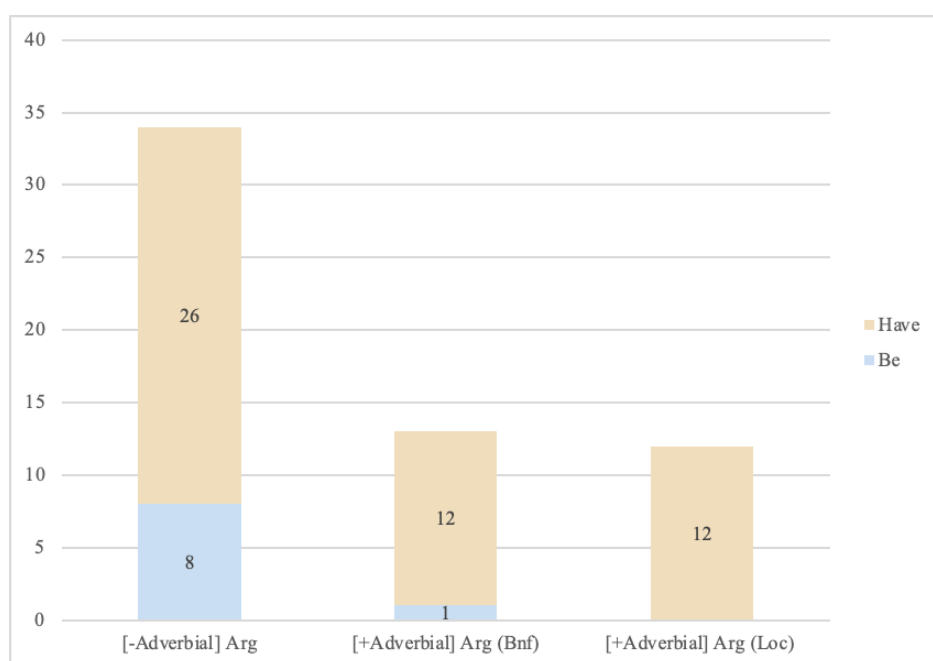


FIGURE 3.10 – Correlation of ‘have’ and ‘be’ and adverbials arguments.

Factors	Data ‘Be’	Data ‘Have’	Total
[-Adverbial] Arg	8	26	36
[+Adverbial] Arg (Bnf)	1	12	13
[+Adverbial] Arg (Loc)	/	12	12
<b>Total</b>	9	50	59

TABLE 3.13 – Frequency of ‘have’ and ‘be’ based on adverbials arguments.

As both the figure and the table show, occurrences with argument adverbials nearly always present ‘have’ (with only one exception). This type of adverbial tends to favor a causative interpretation of the subject. Locative adverbial arguments present both agentive (49) and internal cause subjects (50), while beneficiary ones only present internal cause subjects (51):

- (49) Insomma, io ho contato su di lui e non sono  
 in.short I have.PRS.1SG count.PTCP on of him and not be.PRS.1SG  
 rimasto deluso.  
 be.PTCP disappointed

‘In short, I counted on him and was not disappointed.’

- (50) [...] da sempre ogni civiltà ha contato - lo sapesse o  
 since.time every civilization have.PRS.3SG count.PTCP it know.SUBJ.3SG or  
 meno - sulla sua componente femminile.  
 not on.DEF.ART its component female

[...]since time immemorial every civilization has relied whether it knew it or not – on its female component.’

- (51) Quali personalità della non- hanno contato di più  
 which personalities of.DEF.ART nonviolence have- matter- the.most  
 PRS.3PL PTCP  
 per te?  
 for you

‘Which nonviolence personalities have mattered most to you?’

As illustrated in (49), the locative adverbial argument headed by *su* involves intentional behavior. This type of locative adverbial argument is embedded in a construction that resembles an indirect transitive one more than an intransitive one, which could explain why ‘be’ is completely absent.

(50) is a specific scenario, though, because the parenthetical *lo sapesse o meno* ‘whether it knew it or not’ indicates that the process could have been unintended (with a non-agentive but internal cause subject).

Instead, the beneficiary argument illustrated in (51) involves a subject corresponding to an internal cause.

**QUANTITY ADVERBIALS** Quantity adverbials combining with *contare* ‘matter’ indicate the degree of importance of the entity corresponding to the subject. As shown in Figure 3.11 and Table 3.14, this combination is even rather frequent. These adverbials (*contare molto* ‘matter

a lot’ can co-occur with both ‘have’ and ‘be’, but ‘have’ is preferred once again because of the feature of internal cause. Hence, these adverbials do not by themselves influence the auxiliary selection.

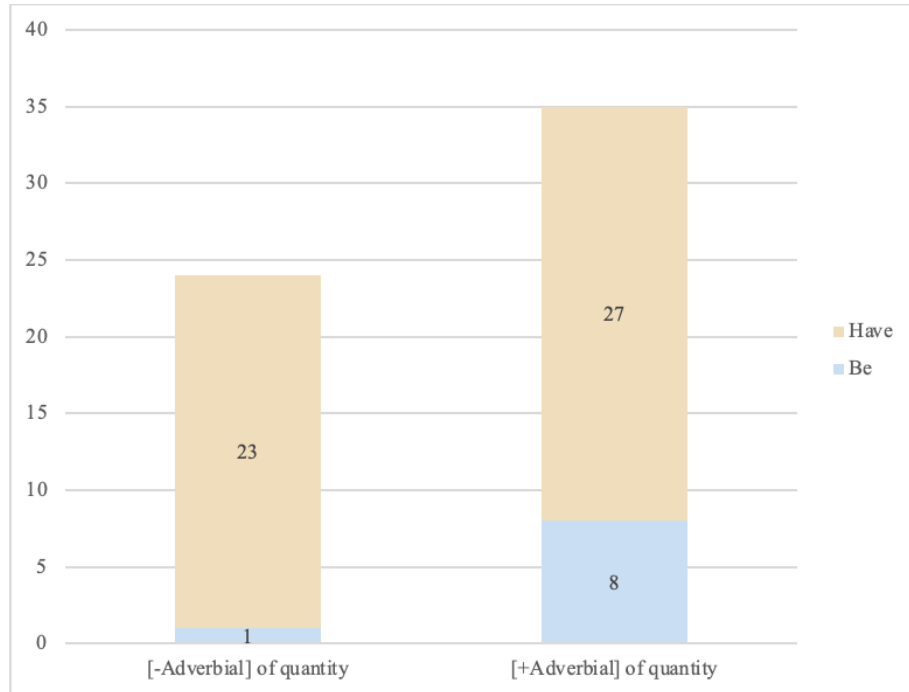


FIGURE 3.11 – Correlation of ‘have’ and ‘be’ and the quantity adverbial.

Factors	Data ‘Be’	Data ‘Have’	Total
[-Adverbial of quantity]	1	23	24
[+Adverbial of quantity]	8	27	35
<b>Total</b>	9	50	59

TABLE 3.14 – Frequency of ‘have’ and ‘be’ depending on the quantity adverbial.

Interestingly, *qualcosa* functions as a quantity adverbial, indicating the extent to which something, whether human (52) or non-human (53; 54), matters

- (52) Il            popolo    ha                    contato    qualcosa    ed    ha  
DEF.ART    people    have.PRS.3SG    count.PTCP    something.ADV    and    have.PRS.3SG
- detto        la            sua.  
say.PTCP    DEF.ART    theirs

‘The people counted for something and had their say.’

(53) Il contesto storico e la posizione della donna nella  
 DEF.ART context historical and DEF.ART position of. woman in.DEF.ART  
 DEF.ART

società ebraica del tempo avrà contato qualcosa?  
 society jewish of. DEF.ART time have.FUT.3SG count.PTCP something.ADV

‘The historical context and the position of women in Jewish society of the time will have mattered?’

(54) Le vendite immani della prima serie son contate  
 DEF.ART sales immense of. DEF.ART first series be-PRS.3PL count.PTCP.F.PL

qualcosa, a quanto pare.  
 something.ADV apparently

‘The immense sales of the first series counted for something, apparently.’

Despite the prevalence of the auxiliary ‘have’, it is intriguing to note that even when modified by a potentially ambiguous adverbial as such as *qualcosa*, instances with ‘be’ are present.

In summary, the analysis underscores ‘have’ as the primary auxiliary associated with *contare* ‘count’, primarily due to the semantic factors of agentivity or internal causation, irrespective of whether the subject is human or non-human. This observation is depicted in Table 3.15.

Agentivity	Human	Auxiliary
Internal Cause	Human/Non-human →	Have
Agentive	Human →	Have

TABLE 3.15 – Principal factors affecting auxiliary distribution for *contare*.

As illustrated in the table, when the subject exhibits agency, it is inherently human, leading to the selection of the auxiliary ‘have’. Conversely, when the subject embodies an internal cause, it can encompass both human and non-human entities, yet the preferred auxiliary remains ‘have’. However, the analysis also some rare instances with ‘be’ as an auxiliary, notably with non-human subjects.

### 3.4.2 *Pesare*

IfTenTen (16) reveals a higher frequency of occurrences of the auxiliary verb ‘have’ compared to ‘be’ in association with the verb *pesare*, which has two distinct meanings, a literal meaning ‘weigh’ and a meaning close to the verb *contare* used intransitively ‘matter’. Specifically,

applying the part-of-speech filter, there is the 71% of ‘have’ comparing to 29% of ‘be’, with 185 occurrences with ‘have’ and 74 with ‘be’.

This preference for ‘have’ over ‘be’ echoes the pattern observed with the verb *contare*, indicating a strong tendency towards the auxiliary ‘have’ in expressing the concept of importance or significance, as embodied in the verb *pesare*.

**Human and non-human subjects** In comparison with *contare*, with *pesare*, non-human subjects are even more predominant. While the auxiliary ‘have’ remains prevalent, there is also a notable proportion of instances with ‘be’, particularly in association with non-human subjects. This observation is supported by the data depicted in Figure 3.12 and Table 3.16.

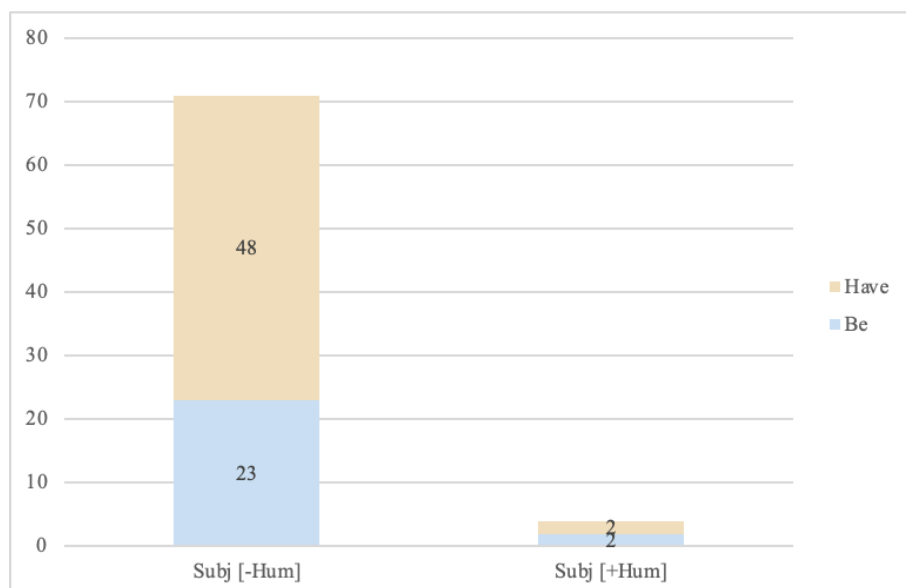


FIGURE 3.12 – Correlation of ‘have’ and ‘be’ with human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	23	48	71
Subj [+Hum]	2	2	4
<b>Total</b>	25	50	75

TABLE 3.16 – Frequency of ‘have’ and ‘be’ based on the human/non-human subjects.

The analysis of sentences involving *pesare* encompasses a total of 75 instances, as there were insufficient data regarding occurrences with ‘be’ that met the study’s criteria. Human subjects occur only four times. Among these, only one instance, viz. example (55), conveys the meaning of ‘matter’. It is noteworthy that this particular instance features a metonymical subject.

(55) Le sue parole di ammirazione hanno pesato molto su  
 DEF.ART his words of admiration have.PRS.3PL weigh.PTCP a.lot on

di me.  
 of me

‘His words of admiration weighed heavily on me. [...]’

In (55), the subject *parole* ‘words’ is considered as human because of metonymy (as could be *sguardo* ‘gaze’). In contrast, when the subject is non-human, *pesare* is typically found with the ‘matter’ meaning and ‘have’ prevails (56):

(56) Ma il ritardo e il mancato confronto di questi mesi  
 but DEF.ART delay and DEF.ART lack.PTCP confrontation of these months

ha pesato [...].  
 have.PRS.3SG weigh.PTCP

‘But the delay and lack of confrontation in recent months has weighed [...].’

**Internal cause subjects** With *pesare*, aside from a few instances that involve non-agentive subjects, nearly all occurrences feature an internal cause subject, as in (55) und (56). This factor is, in principle, important for peripheral verbs. As a matter of fact, the potential ambiguity associated with internal cause—between either absence or presence of agency—is correlated with auxiliary alternation.

However, there is a tendency towards a preference for ‘have’, as indicated in Figure 3.13 and Table 3.17.

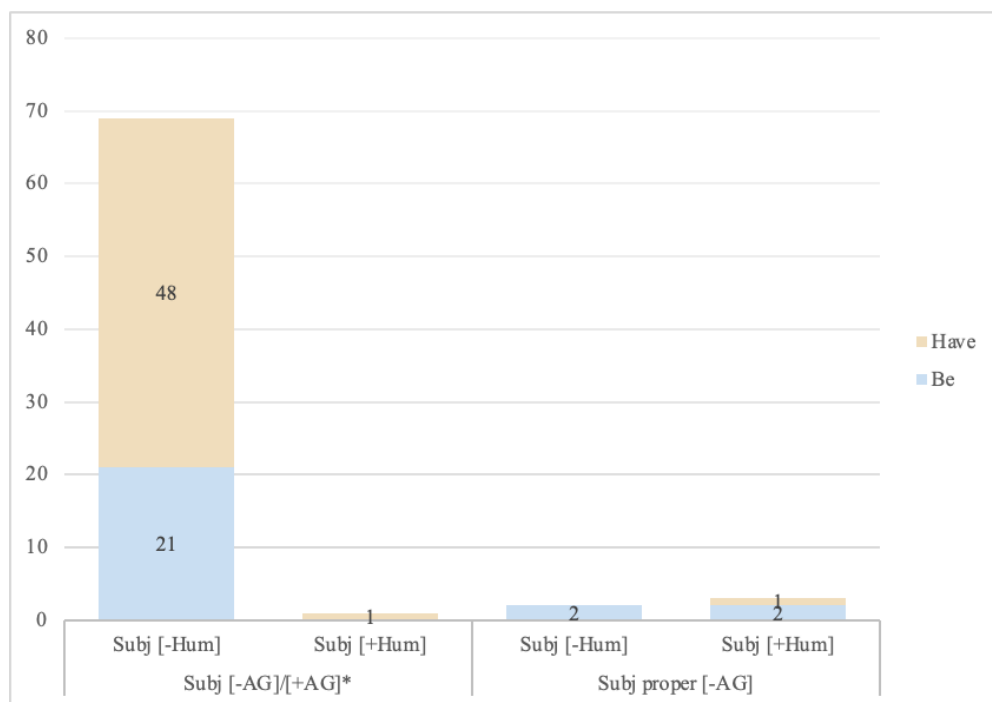


FIGURE 3.13 – Correlation of ‘have’ and ‘be’ with internal cause/non-agentive subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-AG]/[+AG] *[-Hum]</b>	21	48	69
<b>Subj [-AG]/[+AG] * [+Hum]</b>	/	1	1
<b>Subj proper [-AG] [-Hum]</b>	2	/	2
<b>Subj proper [-AG] [+Hum]</b>	2	1	3
<b>Total</b>	25	50	75

TABLE 3.17 – Frequency of ‘have’ and ‘be’ based on internal cause/non-agentive subjects.

Occurrences featuring non-agentive subjects are relatively infrequent. In these cases, the auxiliary ‘be’ tends to predominate, particularly in contexts where the verb’s literal meaning of ‘weigh’ is employed. Consider examples of non-agentive subjects, both human (57) and non-human (58):

- (57) All’arrivo a casa sarò pesato almeno 5 kg in più.  
 at. DEF.ART.arrival at home be.FUT.1SG weigh-PTCP at.least 5 kg more

‘Upon arrival home I will have gained at least 5 kg.’

- (58) I pantaloni saranno pesati almeno 7 kg.  
 DEF.ART pants be.FUT.3PL weigh.PTCP.M.PL at.least 7 kg

‘The pants will have weighed at least 7kg.’

Consequently, non-agentivity co-occurs with the literal sense of ‘weigh,’ which tends to favor the auxiliary ‘be.’ Conversely, when the subject lacks agency but is causative, conveying the meaning of ‘matter,’ there exists a notable alternation between the auxiliaries ‘be’ and ‘have’—with 21 occurrences of ‘be’ compared to 48 occurrences of ‘have’ for non-human subjects.

Although ‘have’ predominates, examples such as (59) and (60) illustrate how both auxiliaries can be employed within the same context:

- (59) Ma un’assenza è pesata più di altre [...].  
 But INDF.ART.absence be.PRS.3SG matter.PTCP.F.SG more than others

‘But one absence mattered more than others.’

- (60) Certo, l’assenza di questa autorità ha pesato.  
 certainly DEF.ART.absence of this authority have.PRS.3SG matter.PTCP

‘Certainly, the absence of this authority has mattered.’

(59) and (60) illustrates the auxiliary selection with the verb *pesare*. The auxiliary ‘have’ is associated with an emphasis on the causative behavior of the subject, highlighting its active role and increasing its impact. Conversely, the auxiliary ‘be’ emphasizes a reading wherein the entity is merely affected, potentially indicating a lack of agency or a passive role.

This selection of auxiliaries may occur unconsciously in the speaker before verbalization, or it could be a deliberate choice to emphasize a particular interpretation.

Another complementary interpretation is that the auxiliary ‘be’ implies a state or a more completed action, making it preferable when describing a concluded event.

**Adverbials** The adverbials associated with the verb *pesare* do not independently influence the selection of the auxiliary verb. However, they serve to reinforce a specific interpretation based on the semantic features of the subject involved. The most frequently used adverbials are adverbial arguments and quantity adverbials.

**ADVERBIAL ARGUMENTS** Two types of locative arguments (*nella sconfitta ha pesato* ‘in the defeat weighed’ and *sulla sentenza ha pesato* ‘on the judgment weighed’ and a dative argument (*a me è pesato* ‘to me has weighed’ occur. With locative arguments, the auxiliary ‘have’ predominates, consistent with the general trend associated with this verb. Conversely, with dative arguments, the auxiliary ‘be’ prevails, as depicted in Figure 3.14 and Table 3.18.

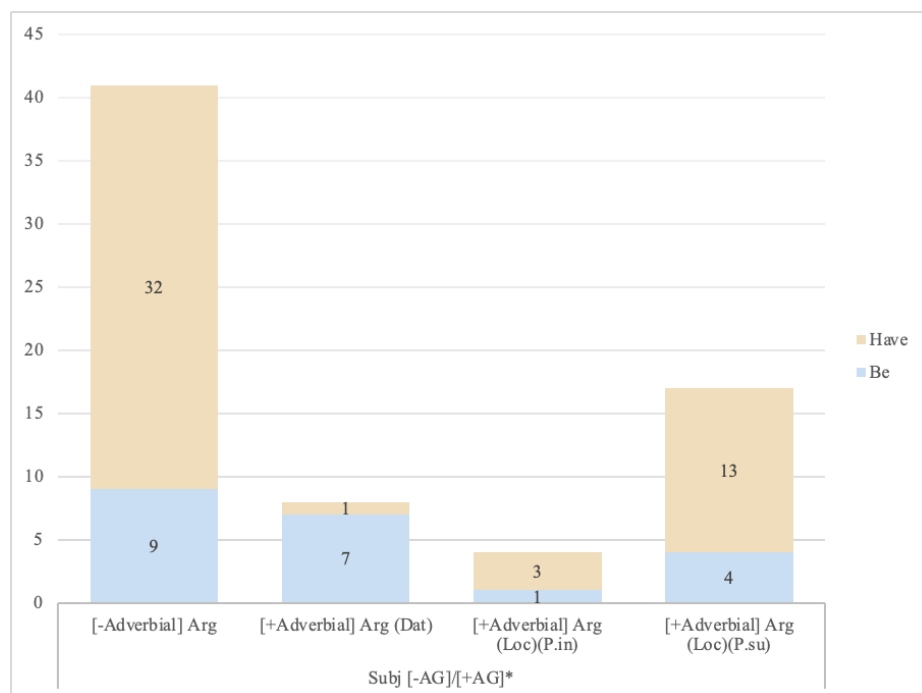


FIGURE 3.14 – Correlation of ‘have’ and ‘be’ with internal cause/non-agentive subjects and adverbial arguments.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG]/[+AG] *[-Adverbial Arg]	9	32	41
Subj [-AG]/[+AG] *[+Adverbial Arg] (Dat)	7	1	8
Subj [-AG]/[+AG] *[+Adverbial Arg] (Loc)(P.in)	1	3	4
Subj [-AG]/[+AG] *[+Adverbial Arg] (Loc)(P.su)	4	13	17
<b>Total</b>	<b>21</b>	<b>49</b>	<b>70</b>

TABLE 3.18 – Frequency of ‘have’ and ‘be’ based on internal cause/non-agentive subjects and adverbial arguments.

As both the figure and table show, even if *pesare* generally can occur with both auxiliaries, ‘be’ prevails with the dative argument (61):

- (61) Sicuramente non avrei ripetuto su loro quello che a me  
 certainly not have.COND.1SG repeat.PTCP on them what that to me  
 è pesato.  
 be.PRS.3SG weigh.PTCP

‘I certainly would not have repeated on them what weighed on me.’

Conversely, with the locative argument, ‘have’ prevails. This is evident in both the locative argument employing *in* (62) and *su* (63). It becomes apparent that with this type of argument,

the emphasis is placed on the subject's impact on the action of *pesare*, thus explaining the predominance of 'have'.

Similarly to *contare*, the use with the locative adverbial argument can be an indirect transitive use.

- (62) Quale mezzo di informazione ha pesato di più nel  
 which medium of information have.PRS.3SG weigh.PTCP more in. DEF.ART  
 decidere?  
 decide.INF

'Which information medium weighed most heavily in the decision?'

- (63) Le donne sono quelle su cui la crisi ha  
 DEF.ART women be.PRS.3PL those on whom DEF.ART crisis have.PRS.3SG  
 pesato e continua a pesare di più.  
 weigh.PTCP and continue.PRS.3SG to weigh.INF more

'Women are the ones on whom the crisis has weighed and continues to weigh most heavily.'

These data provide significant insights: when the prepositions *in* or *su* (on) are used, occurrences with 'have' outnumber those with 'be', suggesting the subject as an entity that imposes weight on or in something else, acting as an internal cause.

Conversely, with a dative argument as in (61), the interpretation of the subject as causative is less pronounced.

**QUANTITY ADVERBIALS** The quantity adverbials themselves do not exert direct influence, but the choice of auxiliary verb accompanying them depends on whether the subject is non-agentive or characterized by internal cause. When the subject lacks agency, the meaning tends to be literal, and 'be' prevails.

Conversely, if the subject is characterized by internal cause, 'have' is predominant, and the adverbials indicate the emotional weight of something on someone. This distribution of 'have' and 'be' with and without quantity adverbials is illustrated in Figure 3.15 and Table 3.19.

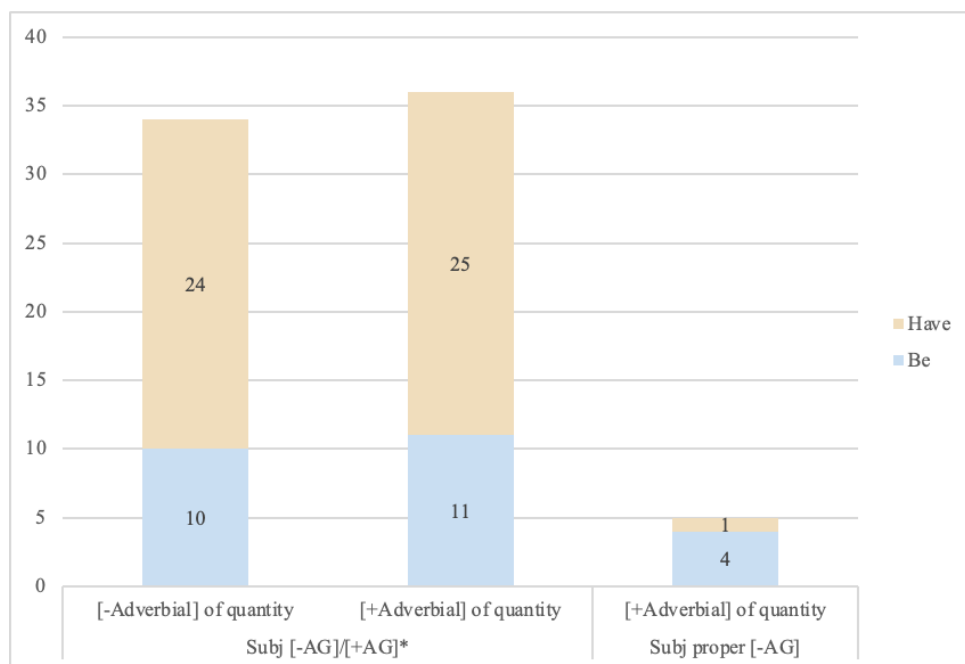


FIGURE 3.15 – Correlation of ‘have’ and ‘be’ with internal cause/non-agentive subjects and quantity adverbials.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG]/[+AG] *[-Adverbial] of quantity	10	24	34
Subj [-AG]/[+AG] *[+Adverbial] of quantity	11	25	36
Subj proper [-AG] [+Adverbial] of quantity	4	1	5
<b>Total</b>	25	50	75

TABLE 3.19 – Frequency of ‘have’ and ‘be’ based on internal cause/non-agentive subjects and quantity adverbials.

In the analysis conducted in this study, *pesare* predominantly appears with internal cause subjects when accompanied by a quantity adverbial (as depicted in Figure 3.15 and Table 3.19). In such instances, ‘have’ emerges as the preferred auxiliary, as exemplified in (64; 65).

Conversely, when the verb is used with non-agentive subjects and a quantity adverbial (*10 tonnellate*), ‘be’ is the preferred auxiliary, as illustrated in (66).

- (64) Quell’eliminazione ha pesato tanto, c’erano tante  
 That.elimitation Have.PRS.3SG Weigh.PTCP much LOC.be.PRS.3PL many

aspettative [...].  
 expectations

‘That elimination weighed so much, there were so many expectations [...].’

(65) In questa edizione il televoto ha pesato per il  
 in this edition DEF.ART televoting have.PRS.3SG weigh.PTCP for DEF.ART

40% sulla classifica finale [...].  
 40% on.DEF.ART ranking Final

‘In this edition, televoting weighed 40% on the final ranking [...].’

(66) vado in ospedale e trovo [...] un ostetrica o assistente  
 go.PRS.1SG to hospital and find.PRS.1SG INDF.ART obstetrician or assistant

dalla delicatezza di un rinoceronte, 10 tonnellate  
 from.DEF.ART delicacy of INDF.ART rhinoceros 10 tons

sarà pesata.  
 be-FUT.3SG weigh-PTCP.F.SG

I go to the hospital and find [...] an obstetrician or assistant from the delicacy of a rhinoceros, 10 tons will be weighed.’

In (64) and (65), two meanings of *pesare* emerge in correspondence with the internal cause of the subject: the emotional meaning of *pesare* as moral influence in (64) and as having influence in (65). Conversely, in (66), the subject is non-agentive, and the meaning conveyed by the verb in the absence of the subject’s agency is the literal meaning of *pesare*.

In conclusion, *pesare* confirms its status as a verb predominantly associated with the auxiliary ‘have’, primarily due to the presence of the internal cause factor.

However, the intriguing nature of internal cause, between agency and lack of agency, accounts for the coexistence of both auxiliaries. Conversely, ‘be’ prevails when the subject lacks agency and consequently does not convey the derived meaning but rather the literal meaning.

Table 3.20 summarizes these key factors.

Agentivity	Human	Auxiliary
Internal Cause	Non-human →	Have (/Be)
Non-agentive	Human/Non-human →	Be

TABLE 3.20 – Most important factors influencing the auxiliary distribution for *pesare*.

### 3.5 VERBS OF CONTIGUOUS LOCATION : *prevalere*

*Prevalere* ‘prevail’ can be considered a verb conveying an existential meaning, as found in Levin (1993: 249). However, if we look at its meaning of predominance or widespread occurrence, its meaning is similar to ‘dominate’, which fits the category of Verbs of Contiguous Location (Levin 1993: 257).

In general, *prevalere* can be considered as a verb of dominance and superiority, where the subject exerts control or influence over others. In the ItTenTen (16) corpus, this verb exhibits both ‘have’ and ‘be’ constructions, although occurrences with ‘have’ are more frequent. Specifically, applying the part-of-speech filter, there is 87% with ‘have’ and 13% with ‘be’, with 401 ‘have’ occurrences and 59 ‘be’ occurrences.

***Human and non-human subjects*** In the sentences analyzed within this study, non-human subjects are particularly prevalent. When the subject is human, the auxiliary ‘have’ overwhelmingly dominates. However, in cases where the subjects are non-human, there is more alternation of auxiliaries, as illustrated in Figure 3.16 and Table 3.21.

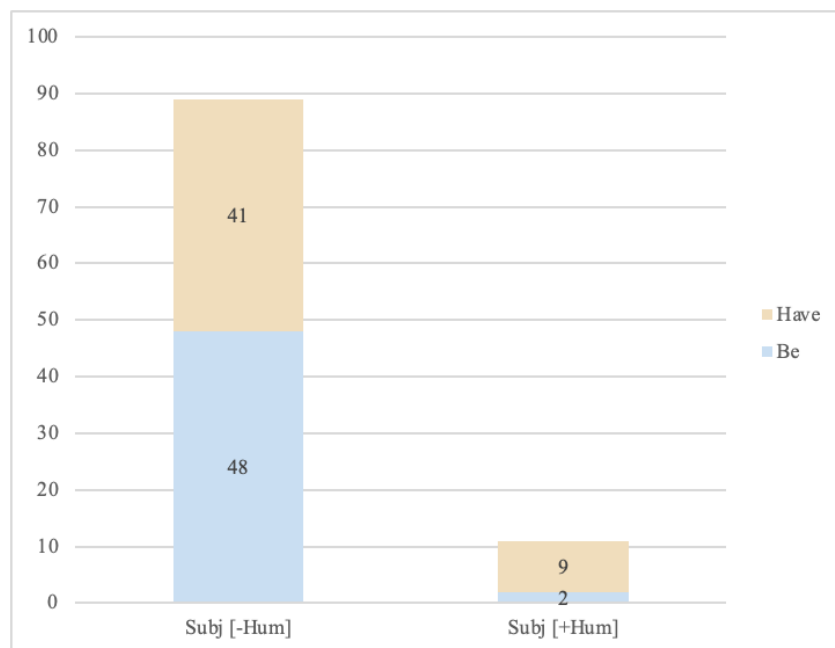


FIGURE 3.16 – Correlation of ‘have’ and ‘be’ with human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	48	41	90
Subj [+Hum]	2	9	10
<b>Total</b>	50	50	100

TABLE 3.21 – Frequency of ‘have’ and ‘be’ depending on the human/non-human nature of the subject.

Consider examples of a human subject with ‘have’ (67) and of non-human subjects with ‘be’ (68) and ‘have’ (69):

- (67) Alla fine ha prevalso solamente chi ha commesso qualche errore in meno dell’avversario.  
 in.DEF.ART end have.PRS.3SG prevail.PTCP only who have.PRS.3SG  
 make.PTCP some mistake in less of.DEF.ART.opponent

‘In the end, only those who made fewer mistakes than their opponent prevailed.’

- (68) Nonostante le ipotesi scaturite sono state diverse, nessuna  
 although DEF.ART hypotheses arise.PTCP be.PRS.3PL be.PTCP several none  
 in particolare è prevalsa.  
 in.particular be.PRS.3SG prevail.PTCP.F.SG

‘Although there were several hypotheses that arose, none in particular prevailed.’

- (69) Alla fine ha prevalso di gran lunga quest’ultima ipotesi.  
 in.DEF.ART end have.PRS.3SG prevail.PTCP by.far this-latter hypothesis

‘The latter hypothesis prevailed by far in the end.’

In (68) and (69), in spite of the same lexical subject, the verb prevail allows both ‘have’ and ‘be’ as an auxiliary. The ability to use both ‘have’ and ‘be’ is attributed to the subject being non-human and internal cause. When the subject is non-human, both interpretations—(i) the subject is an internal cause or (ii) lacks agentivity—are possible. When the subject is human, the internal cause interpretation is more plausible, favouring ‘have’.

**Internal cause** The subject of *prevalere* inherently conveys the feature of internal causation due to the semantics of the verb. In other words, *prevail* implies a causative role of the subject, but no intentionality: the entity corresponding to the subject cannot actively decide to prevail; instead, the process of prevailing happens naturally.

In (70), the causative role of the subject is particularly emphasized:

- (70) Alla sera il sonno ha prevalso su tutti noi,  
 in-DEF.ART evening DEF.ART sleep have.PRS.3SG prevail.PTCP over all us  
 impedendoci di valorizzare anche il momento del documentario.  
 preventing.us of value.INF even DEF.ART moment of.DEF.ART documentary

‘In the evening sleep prevailed over all of us, preventing us from valuing even the documentary moment.’

In (70), the subject *sonno* ‘sleep’ appears to be capable of affecting people, and hence acts as an internal cause. The use of ‘have’ in this context emphasizes the causative action of the subject.

Considering the data presented in Figure 3.16 and Table 3.21, along with the discussion on internal cause, it appears that *prevalere* ‘prevail’ allows for considerable variation with inanimate subjects. This suggests that it should be placed closer to the middle of the hierarchy proposed by Sorace (2000), rather than at the bottom where verbs primarily associated with ‘have’, and thus agentivity, are located. Sorace (2000) admits that ‘be’ can indeed be used with non-human subjects, but my analysis indicates a more balanced distribution between ‘have’ and ‘be’.

Furthermore, the correlation between internal cause and the lack of animacy of the subject can be considered as a criterion influencing auxiliary selection in this context.

**Adverbials** Among the adverbials occurring with *prevalere*, the most frequent ones are adverbial arguments and locative adverbials indicating a static place.

**ADVERBIAL ARGUMENTS** *Prevalere* commonly combines with a locative complement invariably headed by *su* ‘prevail over’, and consequently this complement functions as an adverbial argument. Although the occurrences are not numerous overall, there is a slight predominance of ‘have’, as illustrated in Figure 3.17 and Table 3.22. This suggests that although both auxiliaries are equally admitted, in contexts where the verb denotes prevailing over something, ‘have’ is the preferred auxiliary.

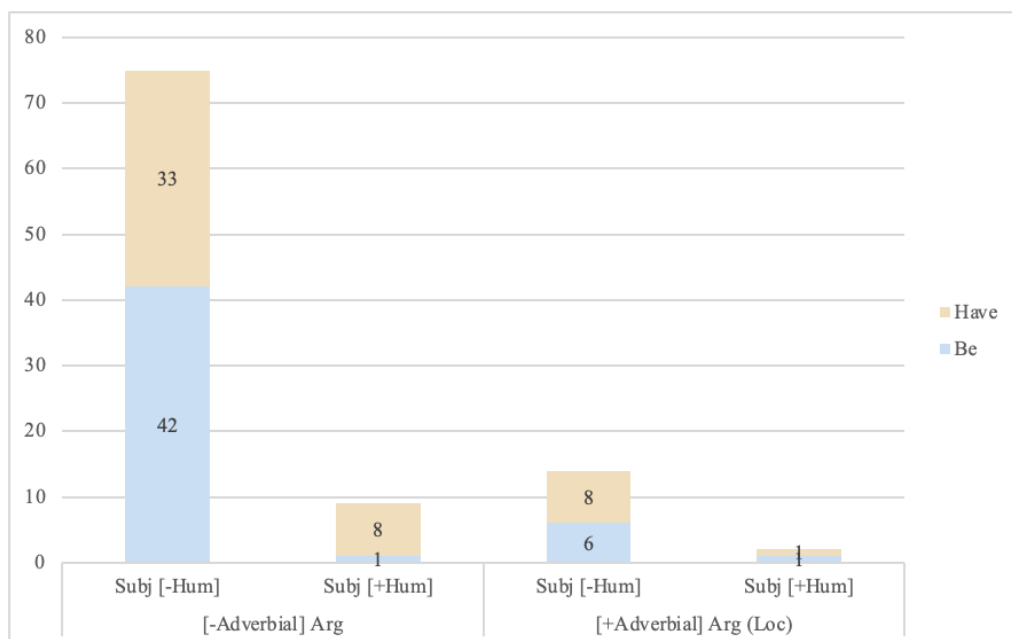


FIGURE 3.17 – Correlation of ‘have’ and ‘be’ with adverbial arguments and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
[-Adverbial] Arg Subj [-Hum]	43	33	76
[-Adverbial] Arg Subj [+Hum]	1	8	8
[+Adverbial] Arg (Loc) Subj [-Hum]	6	8	14
[+Adverbial] Arg (Loc) Subj [+Hum]	1	1	2
<b>Total</b>	50	50	100

TABLE 3.22 – Frequency of ‘have’ and ‘be’ according to adverbial arguments and human/non-human subjects.

Undoubtedly, the usage of ‘be’ is possible and has been observed. The prevalence of ‘have’ over ‘be’ substantiates my earlier observations. Indeed, the expression ‘prevail over someone’ accentuates a sense of dominance exerted by the prevailing entity, leading to considering the subject as an internal cause. Consider example (71):

(71) [...] la natura avrebbe prevalso su tutto ciò che la  
DEF.ART nature have.COND.3SG prevail.PTCP over all it that DEF.ART

civiltà ha costruito.  
civilization have.PRS.3SG build.PTCP

‘[...] nature would prevail over all that civilization has built.’

**LOCATIVE NON-ARGUMENTAL ADVERBIALS** If locative adverbials indicating a static place are considered, such as *in Europa* ‘in Europe’), ‘be’ slightly prevails with non-human subjects,

as illustrated in the Figure 3.18 and Table 3.23. It should be noted, however, that the number of occurrences is small.

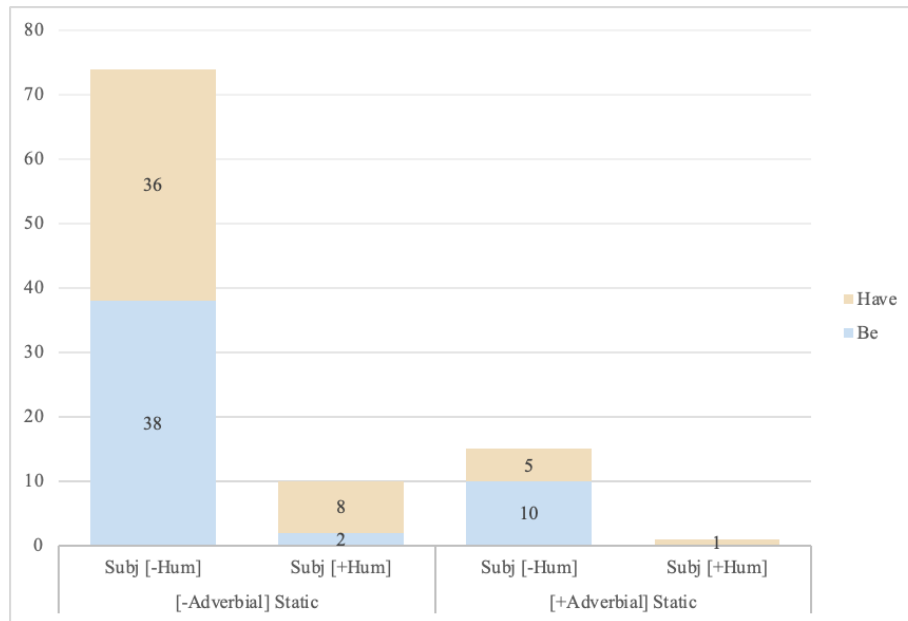


FIGURE 3.18 – Correlation of ‘have’ and ‘be’ with locative adverbials and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
[-Adverbial Static] Subj [-Hum]	38	36	74
[-Adverbial Static] Subj [+Hum]	2	8	10
[+Adverbial Static] Subj [-Hum]	10	5	15
[+Adverbial Static] Subj [+Hum]	/	1	1
<b>Total</b>	50	50	100

TABLE 3.23 – Frequency of ‘have’ and ‘be’ according to locative static adverbials and human/non-human subjects.

As illustrated in the figure and table, both ‘have’ and ‘be’ are employed in this context, although the number of occurrences with the locative adverbial is not high. ‘Have’ is used because the subject of *prevalere* is inherently associated with an internal cause, while ‘be’ is preferred due to the locative nature of the adverbial, which emphasizes a stative interpretation in (72) and a resulting state in (73):

- (72) Nelle zone dove questa religione è prevalsa, sono  
in.DEF.ART areas where this religion be.PRS.3SG prevail.PTCP.F.SG be.PRS.3PL  
  
stati conosciuti come Assiri.  
be.PTCP be.PTCP.M.PL as Assyrians

‘In areas where this religion prevailed, they were known as Assyrians.’

(73) [...]alla Electrolux il no è prevalso per il  
 at.DEF.ART E. DEF.ART no.vote be-PRS.3SG prevail-PTCP by DEF.ART

58% [...].  
 58%

‘[...] at Electrolux the no vote prevailed by 58 percent [...].’

On the one hand, as demonstrated by examples (72) and (73), the presence of the locative adverbial emphasizes the stative interpretation rather than a processive interpretation linked to the internal cause nature of the subject. At the same time, the choice of auxiliary itself can accentuate the internal cause reading when ‘have’ is used and attenuate this reading when ‘be’ is employed.

In conclusion, *prevalere* is a verb that inherently features an internal cause subject due to its semantics. When the subject is human, the preferred auxiliary is ‘have’.

However, when the subject is non-human, there is a more balanced alternation between ‘have’ and ‘be’. The primary factors influencing auxiliary distribution are summarized in Table 3.24.

Agentivity	Human	Auxiliary
Internal Cause	Non-human →	Have /Be
Internal Cause	Human →	Have

TABLE 3.24 – Primary factors influencing the auxiliary selection for *prevalere*.

The verb *prevalere* can be characterized as allowing auxiliary alternation without a preferred auxiliary, as both ‘have’ and ‘be’ are equally frequent. This balance in frequency can be attributed to the internal cause trait of the subject.

### 3.6 VERBS OF FAILURE: *fallire*

*Fallire* ‘fail’ is a verb conveying the meaning of not achieving a desired outcome. In the ItTenTen (16) corpus, it appears with ‘have’ 43% of the time (1,002 occurrences) and with ‘be’ 57% of the time (1,3002 occurrences), with part-of-speech included.

*Human and non-human subjects Fallire* can be accompanied by both ‘have’ and ‘be’, as evidenced by occurrences in the ItTenTen (16) corpus. The choice of auxiliary varies depending on the nature of the subject: ‘be’ is preferred when the subject is non-human, whereas ‘have’ is preferred when the subject is human. This balanced distribution is illustrated in Figure 3.19 and Table 3.25.

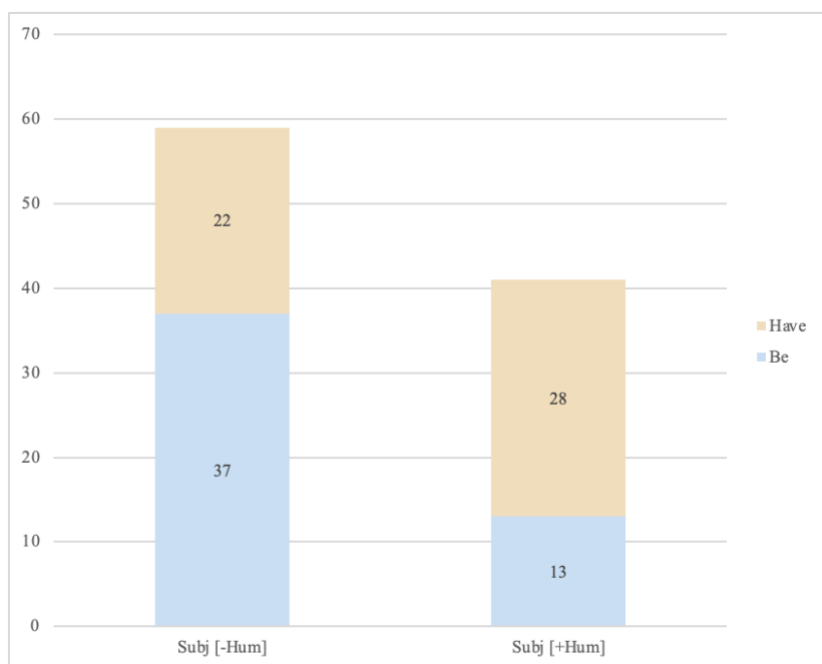


FIGURE 3.19 – Correlation of ‘have’ and ‘be’ with human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	37	22	59
Subj [+Hum]	13	28	41
<b>Total</b>	50	50	100

TABLE 3.25 – Frequency of ‘have’ and ‘be’ depending on human/non-human subjects.

When the subject is human, the preferred auxiliary is ‘have,’ and the conveyed meaning is ‘to fail at something,’ implying a sense of responsibility, as exemplified in (74):

- (74) Anche sul fronte dell’innovazione il governo ha  
 even on.DEF.ART front of.DEF.ART.innovation DEF.ART government have-PRS.3SG  
 fallito [...].  
 fail.PTCP

‘Even on the innovation front, the government has failed.’

Conversely, when the subject is non-human, the preferred auxiliary is ‘be’, and the meaning can encompass both ‘fail at something’ (75) and a resultative meaning of ‘go bankrupt’ (76):

- (75) Non si può dire che l’esperimento sia fallito,  
 not REFL can.PRS.3SG say.INF that DEF.ART.experiment be.SUBJ.3SG fail.PTCP
- ma neanche che abbia sfruttato appieno le  
 but neither that have.SUBJ.3SG take.advantage.PTCP fully DEF.ART
- grandi possibilità.  
 great possibilities

‘It cannot be said that the experiment failed, but neither can it be said that it took full advantage of the great possibilities [...].’

- (76) La sua vita è un disastro; la sua pasticceria  
 DEF.ART his life be.PRS.3SG INDF.ART disaster DEF.ART his bakery
- è fallita [...].  
 be.PRS.3SG fail.PTCP.F.SG

‘His life is a disaster; his bakery is bankrupt.’

While the only interpretation of (76) is that of ‘go bankrupt’, (75) can in principle be interpreted both with this resultative meaning and as ‘fail at something’. However, the latter interpretation is more appropriate due to the subsequent phrase *che abbia sfruttato appieno le grandi possibilità* ‘that it took full advantage of the great possibilities’). This sentence reinforces the interpretation in which the subject plays an active role rather than being a passive participant, indicating an internal cause nature. The reason both interpretations could initially be considered is due to the internal cause nature of the subject, which is expressed only with ‘have’ with human subjects, but with both ‘have’ and ‘be’ with non-human subjects.

**Internal cause and (lack of) agentivity** As shown in Figure 3.20 and Table 3.26, internal cause is a crucial factor for auxiliary selection with *fallire*. When the subject lacks agency, both human and non-human subjects exclusively take ‘be’. Conversely, when the subject is internal cause, two scenarios emerge: human subjects exclusively employ ‘have’, while non-human subjects exhibit an equal distribution of ‘have’ and ‘be’.

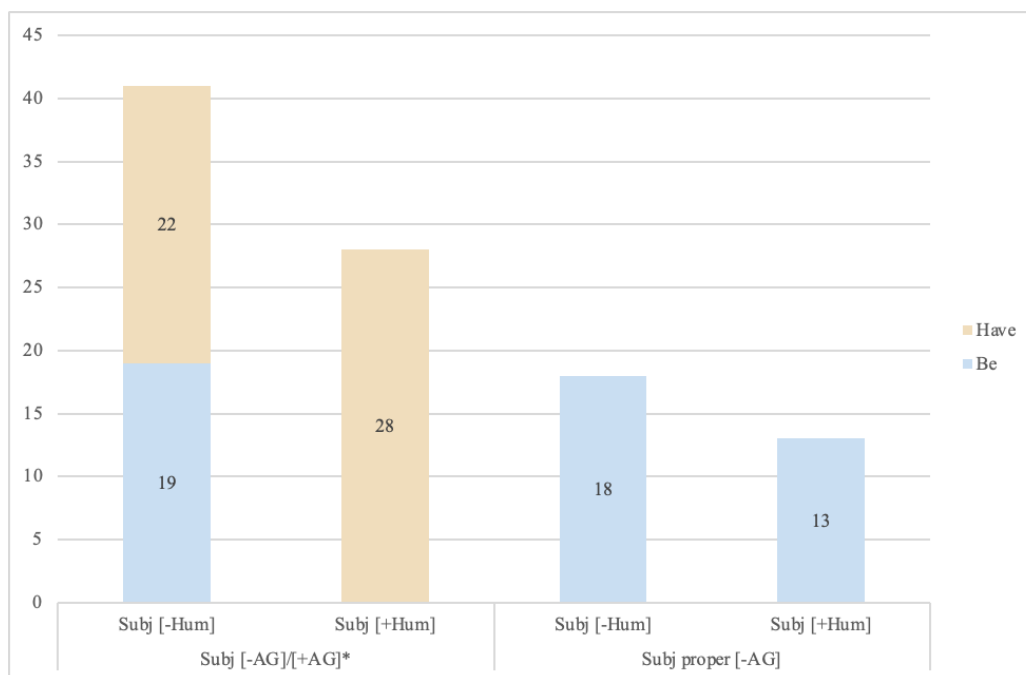


FIGURE 3.20 – Correlation of ‘have’ and ‘be’ with internal cause/lack of agentivity and human/non-human subjects

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG]/[+AG] *[-Hum]	19	22	41
Subj [-AG]/[+AG] *[+Hum]	/	28	28
Subj proper [-AG] [-Hum]	18	/	18
Subj proper [-AG] [+Hum]	13	/	13
<b>Total</b>	<b>50</b>	<b>50</b>	<b>100</b>

TABLE 3.26 – Frequency of ‘have’ and ‘be’ depending on internal cause/lack of agentivity and human/non-human subjects.

Importantly, when the subject is internal cause, *fallire* conveys the meaning of ‘fail at something’: ‘have’ selection with human subjects is illustrated by (77) with human subjects, while the availability of both ‘have and ‘be’ with non-human subjects is illustrated respectively in (78) and (79):

(77) Lo sai perché ogni volta che hai cominciato una  
 it know.PRS.2SG why every time that have.PRS.2SG start.PTCP INDF.ART  
 dieta hai fallito?  
 diet have.PRS.2SG fail.PTCP

‘Do you know why every time you started a diet you failed?’

(78) Tutti i tentativi hanno fallito miseramente.  
 all DEF.ART attempts have.PRS.3PL fail.PTCP miserably

‘All attempts have failed miserably.’

(79) Tutti i miei tentativi sono falliti.  
 All DEF.ART my attempts be.PRS.3PL fail.PTCP.M.PL

‘All my attempts have failed.’

The non-human subject *tentativi* is considered causative due to its strong association with a human trait (an attempt implies a person trying). Moreover, both (78) and (79) convey the notion of something that has failed. With a non-human subject, this can be expressed by means of both ‘have’ and ‘be’.

In contrast, the meaning of ‘go bankrupt’ can only be conveyed through the ‘be’-construction, and in this scenario, the subject is non-agentive. Consider examples (80) and (81):

(80) Tante aziende produttrici di ortaggi sono fallite o  
 so.many farms producers of vegetables be.PRS.3PL fail.PTCP.F.PL or

stanno chiudendo.  
 are.closing.down

‘So many vegetable farms have gone out of business or are closing down.’

(81) Il sistema economico basato sull’indebitamento è fallito  
 DEF.ART system economic base. PTCP on.DEF.ART.debt be.PRS.3SG fail. PTCP

e noi ne pagheremo le conseguenze.  
 and we of.it pay.FUT.3SG DEF.ART consequences

‘The economic system based on debt has failed and we will pay the consequences.’

In (80) and (81), when the subject is non-agentive, the verb *fallire* conveys a resultative meaning. In (81), it indicates the outcome or result of the system’s performance or lack thereof. Similarly, in (80), it signifies that the farms did not succeed or achieve their objectives, resulting in a state of failure, as emphasized by the phrase *o stanno chiudendo* ‘or they are closing down’).

**Adverbials** The most frequent adverbial occurring with *fallire* is the adverbial argument *in* ‘in’ or *dove* ‘where’, as seen in examples such as *Facebook è riuscito dove la lotta al terrorismo ha fallito* ‘Facebook has succeeded where counterterrorism has failed’).

All 13 occurrences feature internal cause subjects, as expected, given that the construction *fallire in* implies a causative role of the subject in the action. The only auxiliary used is ‘have’, both with non-human (82) and human subjects (83):

(82) Facebook è riuscito dove la lotta al terrorismo  
 Facebook be.PRS.3SG succeed.PTCP where DEF.ART fight at.DEF.ART terrorism

ha fallito [...].  
 have.PRS.3SG fail.PTCP

‘Facebook has succeeded where the fight against terrorism has failed [...].’

(83) Vanga riuscì là dove il suo predecessore aveva  
 V. succeed.PST.3SG there where DEF.ART his predecessor have.IPFV.3SG

fallito.  
 fail.PTCP

‘Vanga succeeded where his predecessor had failed.’

The expression *riuscire dove x ha fallito* ‘succeed where x has failed’) appears in both sentences. The adverb *dove* ‘where’ indicates the context or situation in which the action occurred in the subordinate clause. While the adverbial itself does not influence the auxiliary distribution, it underscores the internal cause nature of the subject.

To conclude, in the case of *fallire*, the primary factor explaining the distribution between ‘have’ and ‘be’ is internal cause: when the subject lacks agency, the auxiliary is ‘be’.

Otherwise, it is ‘have’ or ‘be’ depending on the nature of the subject. Specifically, when the subject is human, the auxiliary is ‘have’, whereas when the subject is non-human, both auxiliaries are possible.

Table 3.27 provides an overview of these factors.

Agentivity	Human	Auxiliary
Internal Cause	Non-human →	Have /Be
Internal Cause	Human →	Have
Non-agentive	Human/Non-human →	Be

TABLE 3.27 – Primary factors influencing the auxiliary selection for *fallire*.

As shown in Table 3.27, the choice of auxiliaries is determined by an interplay of multiple factors, specifically internal cause and the human or non-human nature of the subject. When the subject is internal cause, emphasizing the causative nature of the subject, ‘have’ prevails and is even the only option with human subjects. Conversely, when the subject is non-agentive, indicating a resultative meaning, the auxiliary is ‘be’.

### 3.7 VERBS INDICATING A STAGE OF A PROCESS

These verbs encompass various stages or aspects of a dynamic situation or a process, ranging from initiating or commencing this process (as seen in verbs like *iniziare* and *cominciare* ‘start’ ‘begin’), progressing (as in *procedere* ‘proceed’), or continuing it (such as *continuare* and *proseguire* ‘go on’ ‘continue’), or completing it (as exemplified by *finire* ‘finish’). All these verbs are labile verbs as they can be used transitively and intransitively with the same meaning (e.g., *Ho cominciato una nuova avventura* – ‘I have begun a new adventure’ and *Una nuova avventura è cominciata* – ‘A new adventure began’).

We will consider in this section only their use as a full lexical verb, postponing the examination of their use as a semi-auxiliary to section 3.9.

#### 3.7.1 VERBS INDICATING THE START OF A PROCESS: *iniziare*

In ItTenTen (16), *iniziare* ‘begin’ mostly appears with ‘be’ rather than ‘have’. Specifically, the verb appears in 766 occurrences with ‘have’ and in 3,475 occurrences with ‘be’, amounting to 18% of the time with ‘have’ and 82% with ‘be’ when applying a part-of-speech context that excludes nouns within the first five words.

*Iniziare*, when used intransitively, combines mostly with inanimate subjects and the auxiliary ‘be’, like in (84):

(84) La riunione è iniziata intorno alle 10.15.  
 DEF.ART meeting be.PRS.3SG start.PTCP.F.SG around at DEF.ART 10.15 a.m.

‘The meeting started around 10:15 a.m.’

However, in this work, the sometimes fine borderline between transitivity and intransitivity is critical for understanding how potentially predictive elements of auxiliary selection act. As a result, occurrences where the object can be inferred, as shown in (85), are distinguished from the purely intransitive constructions presenting ‘have’, as illustrated in (86):

(85) Mi sono accorto di un merlo che a mezzanotte [...]  
 REFL be.PRS.1SG aware of INDF.ART blackbird that at midnight

cantava a squarciagola [...]. E nei primi giorni di gennaio  
 sing-IPFV.3SG at.loudest.pitch and in DEF.ART first days of january

hanno iniziato anche le cinciallegre.  
 have-PRS.3PL start.PTCP even DEF.ART titmice

‘I noticed a blackbird at midnight [...] singing at the loudest pitch. [...] And in the first days of January, the titmice also started.’

(86) I grandi maghi hanno iniziato non essendo più  
 DEF.ART great magicians have.PRS.3PL start.PTCP not being more

di ciò che siamo noi ora: studenti!  
 than what that be.PRS.3PL we now students

‘The great magicians started out as nothing more than what we are now: students!’

While in (85), an internal object, viz. ‘singing’ could be inferred from the context, in (86), there is no distinct implicit object.

As a result, dividing the usages into those regarded as transitive and intransitive, only one transitive example was identified, namely the one described in (85).

***Human and non-human subjects*** The nature of the subject, whether human or non-human, determines the auxiliary used with this verb. When the subject is non-human, the auxiliary is ‘be’, whereas with a human subject, the auxiliary is ‘have’, as depicted in Figure 3.21 and Table 3.28.

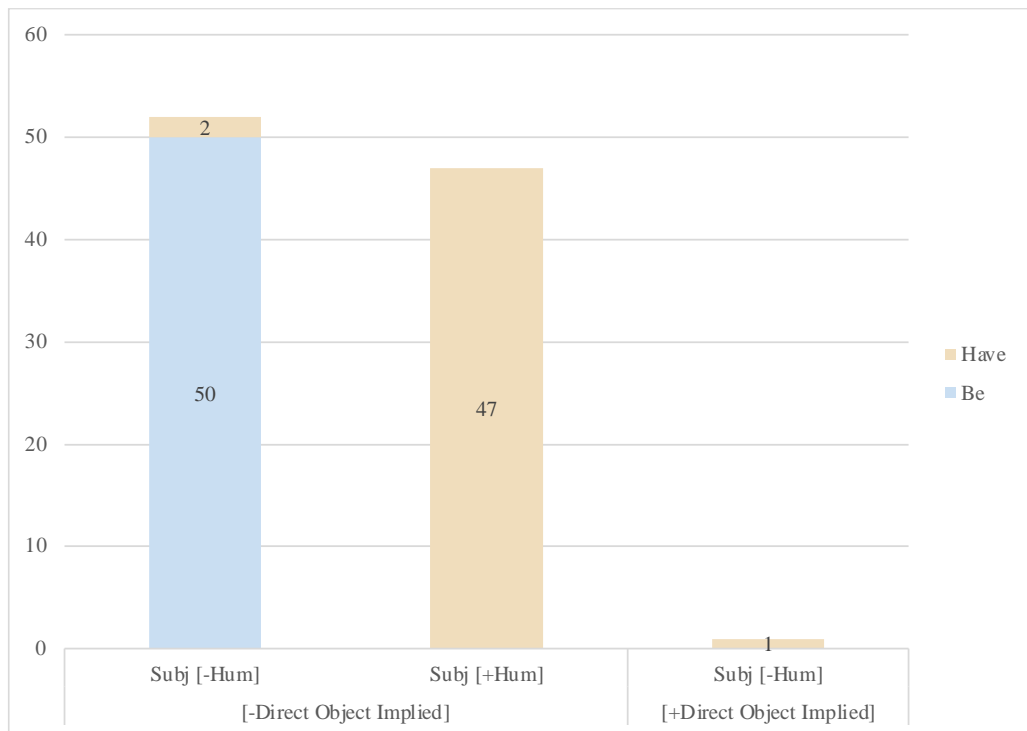


FIGURE 3.21 – Correlation of ‘have’ and ‘be’ with implicit direct object and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] Subj [-Hum]</b>	50	2	52
<b>[-DO] Subj [+Hum]</b>	/	47	47
<b>[+DO] Subj [-Hum] ([+Anim])</b>	/	1	1
<b>Total</b>	50	50	100

TABLE 3.28 - Frequences of ‘have’ and ‘be’ depending on human/non-human subjects.

Figure 3.21 and Table 3.28 primarily focus on intransitive occurrences, with only one transitive occurrence included. Among the intransitives, a clear tendency emerges: ‘be’ is used with non-human subjects (as shown in example 87), while ‘have’ is employed with human subjects (as illustrated in example 88):

(87) La stagione delle piogge è iniziata [...]  
 DEF.ART season of. DEF.ART rains be.PRS.3SG start.PTCP.F.SG

‘The rainy season has begun [...].’

(88) Quando papà ha iniziato, la nostra era solo una  
 When father have.PRS.3SG start.PTCP DEF.ART ours be.IPFV.3SG just INDF.ART  
 libreria.  
 bookstore

‘When father started, ours was just a bookstore.’

The difference does not solely lie in the subject’s nature but also in the agentivity of the subject: in example (87), the subject is non-agentive, whereas in example (88), it is agentive. Specifically, in (88), the subject initiates an action (not specified) with volition. Therefore, the distinction between agentivity, lack of agentivity, and internal cause has to be considered.

**Agentivity, lack of agentivity and internal cause** This semantic nature of the subject, viz. [ $\pm$ HUMAN], is central and explains the auxiliary choice: human subjects occur with ‘have’ when they are agentive or internal cause, while non-human subjects occur with ‘be’ whether they are non-agentive or internal cause. Figure 3.22 and Table 3.29 illustrate this consistent pattern.

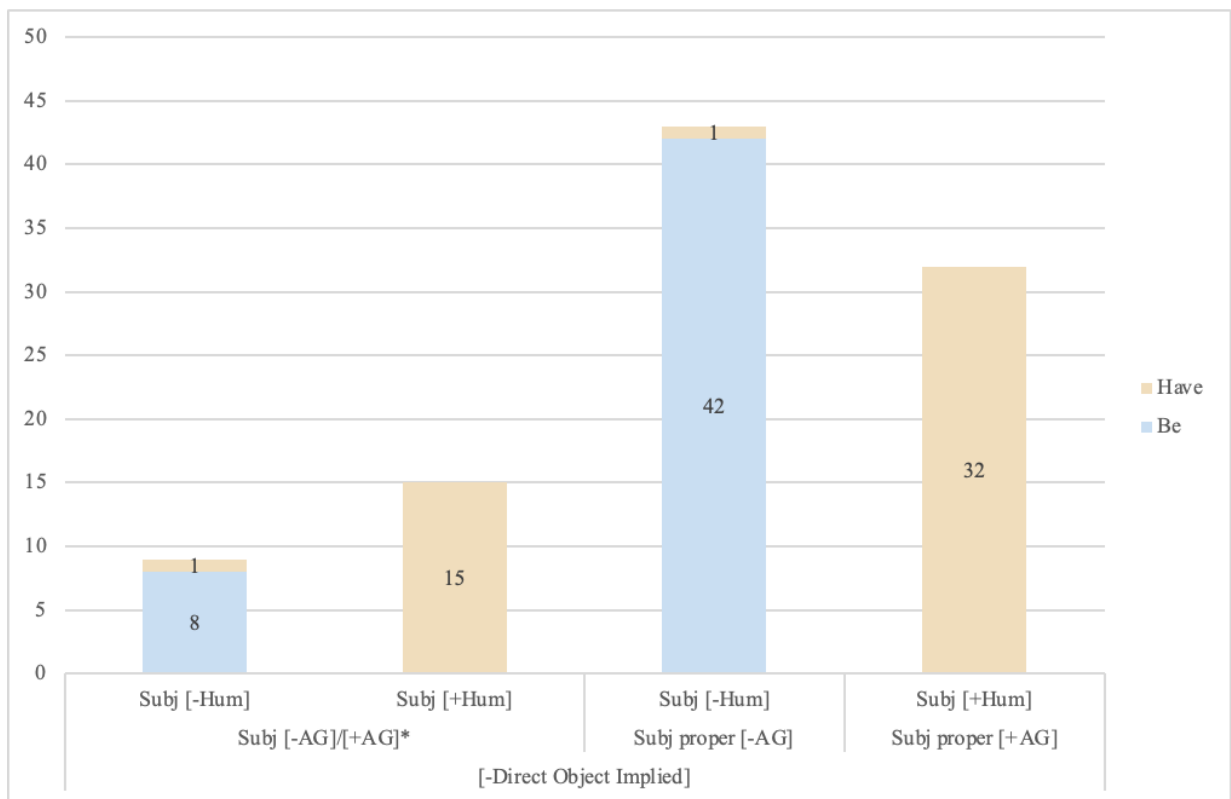


FIGURE 3.22 – Correlation of ‘have’ and ‘be’ with human/non-human and agentive, internal cause and non-agentive subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] Subj [-AG]/[+AG] * [-Hum]</b>	8	<b>1</b>	15
<b>[-DO] Subj [-AG]/[+AG] * [+Hum]</b>	/	<b>15</b>	15
<b>[-DO] Subj proper [-AG] [-Hum]</b>	42	1	43
<b>[-DO] Subj proper [+AG] [+Hum]</b>	/	32	32
<b>Total</b>	50	49	99

TABLE 3.29 - Frequency of ‘have’ and ‘be’ depending on agentive, internal cause and non-agentive subjects and human/non-human subjects.

When the subject of *iniziare* is non-agentive, it is typically non-human, and the auxiliary is ‘be’ (as in example 89), while an agentive subject of *iniziare* must be human and can only select ‘have’ (as illustrated in example 90, where agentivity is highlighted by the manner adverbial *con l’organizzare i viaggi* ‘with organizing trips’):

- (89) La sua carriera era iniziata da giovanissimo [...].  
DEF.ART his career be.IPFV.3SG start.PTCP.F.SG since very.young

‘His career had begun at a very young age.’

- (90) Con questa filosofia abbiamo iniziato con l’organizzare i  
with this philosophy have.PRS.1PL start.PTCP with DEF.ART.organize.INF DEF.ART  
 viaggi di ogni tipo [...].  
trips of all kind

‘With this philosophy we started years ago with organizing trips of all kinds [...].’

**Agentivity, lack of agentivity and internal cause** Similarly, with internal cause subjects, the [ $\pm$ HUMAN] nature of the subject is crucial. The intermediate nature between the lack of agentivity and the causative behavior of internal cause allows non-human subjects to select ‘be’ (as illustrated in example 91) because of the focus on the lack of agentivity, while human subjects select ‘have’ (as demonstrated in example 92) because of the focus on the causativity of the subject.

- (91) [...] la mia mania di loggare è iniziata nel lontano  
DEF.ART my mania of log-INF be-PRS.3SG start.PTCP.F.SG in. DEF.ART far

2002.  
 2002

‘[...] my logging mania started back in 2002.’

- (92) Invece quest’anno avete iniziato alla grande e ne è  
instead this.year have.PRS.2PL start.PTCP greatly and of.it be.PRS.3SG  
 una conferma la vittoria in Coppa Italia.  
INDF.ART confirmation DEF.ART victory in Cup.Italian

‘Instead, this year you’ve gotten off to a great start, and the victory in the Italian Cup is a confirmation of that.’

In (92), the internal-cause nature of the human subject is suggested by the adverb *alla grande*. In fact, the subject cannot decide to start greatly and win the Italian Cup. In (91), the subject *mania* is an internal cause because of its strong link with a human referent.

One of the two exceptions with non-human subjects presents ‘have’ because of the same reason. Consider example (93):

(93) [...] i lavori hanno iniziato nel 1880 [...].  
DEF.ART works have.PRS.3PL start.PTCP in.DEF.ART 1880

‘[...] work began in 1880 [...].’

*Lavori* ‘works’ although unusual as a possibly causative subject can be understood as such because of its strong connection with human action. At the same time, the presence of ‘have’ itself imparts a sense of causativity to the subject. Nevertheless, the auxiliary normally expected for inanimate subjects is ‘be’, as illustrated in example (94):

(94) I lavori sono iniziati nel 2007 [...].  
DEF.ART works be.PRS.3PL start.PTCP.M.PL in.DEF.ART 2007

‘Work began in 2007 [...].’

**Adverbials** Among the most frequent adverbials, there is the temporal location adverbs, e.g. *nel 2007* ‘in 2007’ in (94). However, they do not influence auxiliary selection.

To summarize, with this verb, the relevant factor for auxiliary selection is the human trait as it predominantly co-occurs with agentivity. Additionally, when there is an internal cause, the human nature tends to emphasize the causativity and thus leads to selection of ‘have’.

Conversely, non-human subjects mostly occur with ‘be’ due to the lack of agentivity. When there is internal cause, the non-human nature tends to emphasize the lack of agentivity, leading to the selection of ‘be’.

Table 3.30 summarizes the most influential factors.

Agentivity	Human	Auxiliary
Internal Cause/Agentive	Human →	Have
Internal Cause/Non-agentive	Non-human →	Be

TABLE 3.30 – Primary factors influencing the auxiliary selection for *iniziare*.

### 3.7.2 VERBS INDICATING THE START OF A PROCESS: *cominciare*

*Cominciare* ‘begin’ is more frequently used with ‘be’ than with ‘have’ in the ItTenTen (16) corpus. There are 29% instances of ‘have’ (460 occurrences) compared to 71% occurrences of ‘be’ (1,147) when applying a part of speech filter that excludes nouns in the first five words following the past participle. As with *iniziare*, the intransitive use is predominantly with non-human subjects and ‘be’.

**Human and non-human subjects** Like *iniziare*, *cominicare* is typically used with ‘be’ and non-human subjects. Conversely, when the subject is human, the auxiliary is ‘have’. These tendencies for the intransitive use of *iniziare* are represented in Figure 3.23 and Table 3.31.

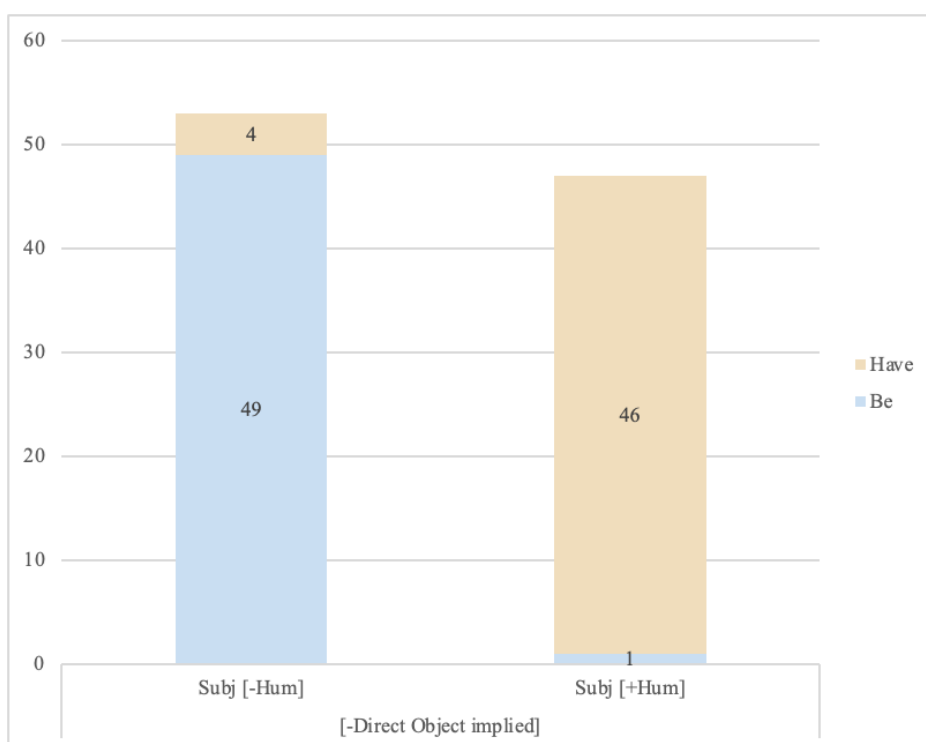


FIGURE 3.23 – Correlation of ‘have’ and ‘be’ with human/non-human subjects excluding direct objects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] Subj [-Hum]</b>	49	4	53
<b>[-DO] Subj [+Hum]</b>	1	46	47
<b>Total</b>	50	50	100

TABLE 3.31 – Frequency of ‘have’ and ‘be’ based on human/non-human subjects excluding direct objects.

Examples of a non-human subject with ‘be’ (95), and a human with ‘have’ are provided in respectively (95) and (96).

(95) Il concerto è cominciato alle 22:30.  
 DEF.ART concert be.PRS.3SG begin.PTCP at. DEF.ART 10:30 p.m.

‘The concert began at 10:30 p.m.’

(96) [...] come scrittrice hai cominciato un po’ tardi.  
 as writer have.PRS.2SG begin.PTCP a.bit late

‘[...] as a writer you began a bit late.’

In sum, the key difference here, as with *iniziare* ‘start’, is whether the subject is human or inhuman. And, as for *iniziare*, this difference is also linked to the agentive or non-agentive nature of the subject.

**Agentivity, lack of agentivity and internal cause** Human subjects correlate with agentivity and internal cause, and in this case, the auxiliary chosen is ‘have’. Conversely, non-human subjects correlate with lack of agentivity and internal cause, and in this case, the auxiliary chosen is ‘be’, as illustrated in Figure 3.24 and Table 3.32.

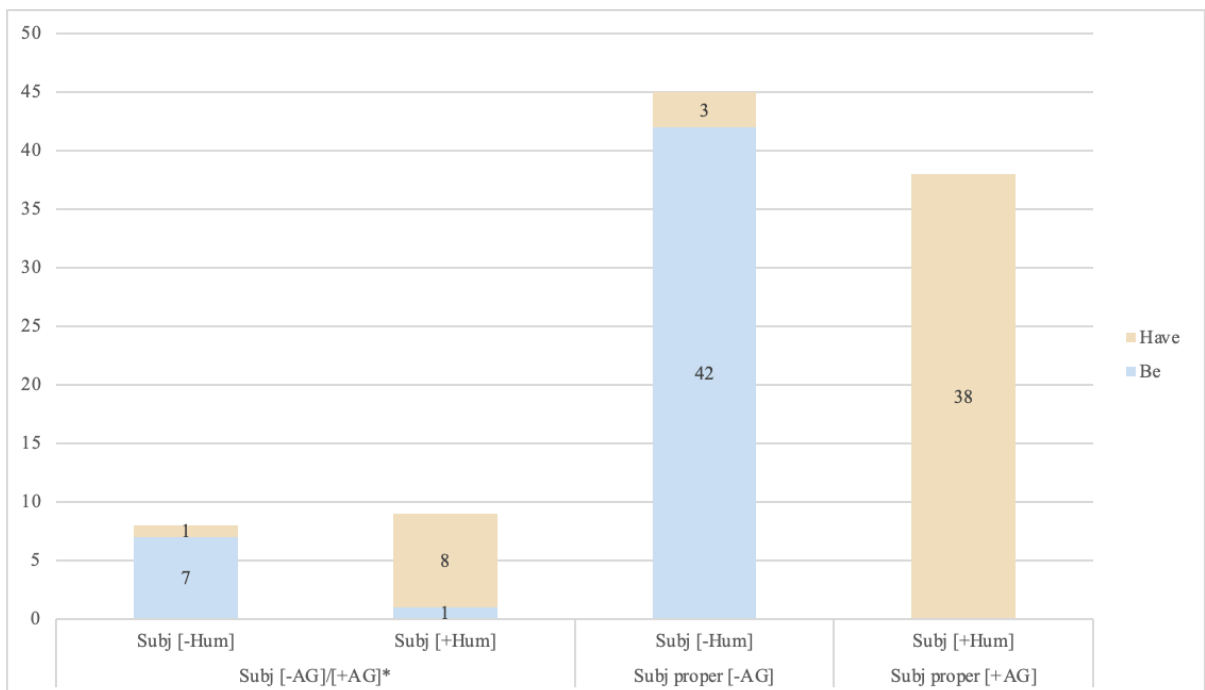


FIGURE 3.24 – Correlation of ‘have’ and ‘be’ with human/non-human and agentive, internal cause and not agentive subjects.

Factors	Data 'Be'	Data 'Have'	Total
Subj [-AG]/[+AG] *[-Hum]	7	1	8
Subj [-AG]/[+AG] *[+Hum]	1	8	9
Subj proper [-AG] [-Hum]	42	3	45
Subj proper [+AG] [+Hum]	/	38	38
<b>Total</b>	50	50	100

TABLE 3.32 – Frequency of ‘have’ and ‘be’ based on human/non-human subjects and agentive, internal cause and non-agentive subjects.

As the Figure shows, human subjects can be completely agentive (97) or non-agentive but causative (98), and in all scenarios, the only auxiliary possible is ‘have’.

- (97) Molti che poi diventano grandi musicisti hanno cominciato  
 many who later become.PRS.3PL great musicians have.PRS.3PL begin.PTCP  
 così.  
 this.way

‘Many who later become great musicians began this way [...]’

- (98) Come attore ho cominciato molto bene, e poi ho  
 As actor have.PRS.1SG begin.PTCP very well and then have.PRS.1SG  
 via via perso quota [...].  
 gradually loose.PTCP level

‘As an actor I started very well, and then gradually lost level [...].’

In (98), the subject is causative as the person cannot decide to begin very well, even if he puts efforts in (and for this reason, he has an active role). As with *iniziare*, the internal cause nature of the subject combined with the human trait of the subject emphasizes the causativity leading to the selection of ‘have’.

Conversely, when the subject is non-human, the internal cause combined with the non-human nature emphasizes the lack of agentivity trait of internal cause (internal cause combines lack of agentivity and causative behavior) and leads to the selection of ‘be’.

Consider the ‘be’ selection with non-human non-agentive subject (99) and non-human internal cause subject, because of the link with a human referent (100):

- (99) [...] i miei problemi di salute sono cominciati nel 2005.  
 DEF.ART my problems of health be-PRS.3PL begin.PTCP.M.PL in.DEF.ART 2005

‘[...]my health problems began in 2005[...].’

(100) La passione per gli impasti è cominciata lì ed  
DEF.ART passion for DEF.ART doughs be.PRS.3SG begin.PTCP.M.SG there and  
 è viva tutt'oggi.  
be.PRS.3SG alive to.this.day

‘The passion for doughs began there and is alive to this day.’

**Adverbials** The most frequent type are time adverbials such as *quest'anno* ‘this year’, with 47 occurrences in total. However, it does not have alter auxiliary selection in function of the major factor: human subjects only select ‘have’, and with non-human subjects, ‘be’ is favored.

Therefore, the human trait and the correlated agentivity/lack of agentivity feature have an influence. This tendency is depicted in Table 3.33

Agentivity	Human	Auxiliary
Internal Cause/Agentive	Human →	Have
Internal Cause/Non-agentive	Non-human →	Be

TABLE 3.33 – Primary factors influencing the auxiliary selection for *cominciare*.

Table 3.33 shows that the primary factors influencing the auxiliary selection for *cominciare*, are identical to that of *iniziare*, with which it is in a relationship of parasynonymy. Auxiliary selection is conditioned by the human or non-human nature of the subject: human subjects select ‘have’ and non-human subjects select ‘be’. Another correlated factor is agentivity: agentivity correlates with human subjects, which tend to select ‘have’, while non-human subjects of *cominciare* are mostly non-agentive and select ‘be’.

Finally, as to internal cause subjects, causativity is amplified for human subjects whereas lack of agentivity is highlighted for non-human subjects, as it resides at the intersection between lack of agentivity and causativity. Additionally, ‘have’ can underscore internal cause when the subject is human.

### 3.7.3 VERBS INDICATING THE PROGRESSION OF A PROCESS: *procedere*

*Procedere* ‘go on’ exhibits a higher frequency with ‘have’ compared to ‘be’. In the ItTenTen (16) corpus, 69% of occurrences are with ‘have’(66) and 31% with ‘be’ (30) when applying the part-of-speech filter.

**Human and non-human subjects** Auxiliary selection for *procedere* is influenced by the nature of the subject, whether human or non-human. Human subjects exclusively select ‘have’, whereas non-human subjects predominantly opt for ‘be’, as illustrated in Figure 3.25 and Table 3.34.

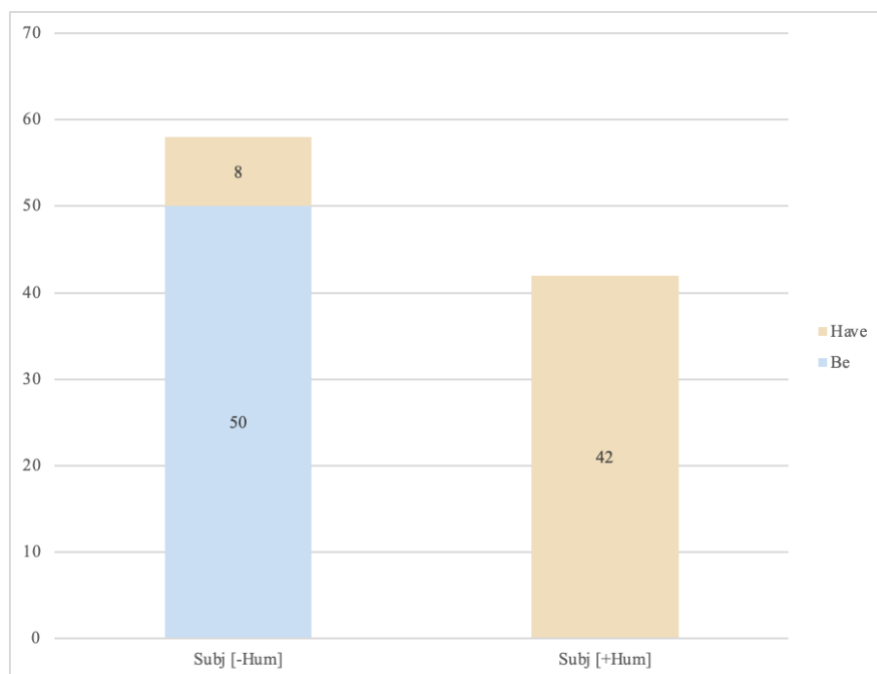


FIGURE 3.25 – Correlation of ‘have’ and ‘be’ with human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-Hum]</b>	50	8	58
<b>Subj [+Hum]</b>	/	42	42
<b>Total</b>	50	50	100

TABLE 3.34 – Frequency of ‘have’ and ‘be’ based on human/non-human subjects.

Examples are provided in (102), with a human subject paired with ‘have’ and (101) containing a non-human subject paired with ‘be’:

(101) La serata è proceduta molto bene [...].  
DEF.ART evening be.PRS.3SG proceed.PTCP.F.SG very well

‘The evening proceeded very well [...].’

(102) Le due cordate avevano proceduto affiancate per non  
DEF.ART two rope have.IPFV.3PL proceed.PTCP side.by.side to not  
teams

danneggiarsi a vicenda.  
damage.REFL each.other

‘The two rope teams had proceeded side by side so as not to damage each other.’

In (102) the subject *cordate* ‘rope teams’ is considered human because it refers to people who are tied together by a rope during a trip. As is the case with *cominciare* and *iniziare*, human subjects mostly exhibit agentivity, while non-human subjects lack it.

**Agentivity, lack of agentivity and internal cause** As previously observed with other dynamic verbs, human subjects are predominantly agentive or internal cause, resulting in the selection of ‘have’. On the other hand, non-human subjects tend to be non-agentive or internal cause, correlating with the auxiliary ‘be’, as illustrated in Figure 3.26 and Table 3.35.

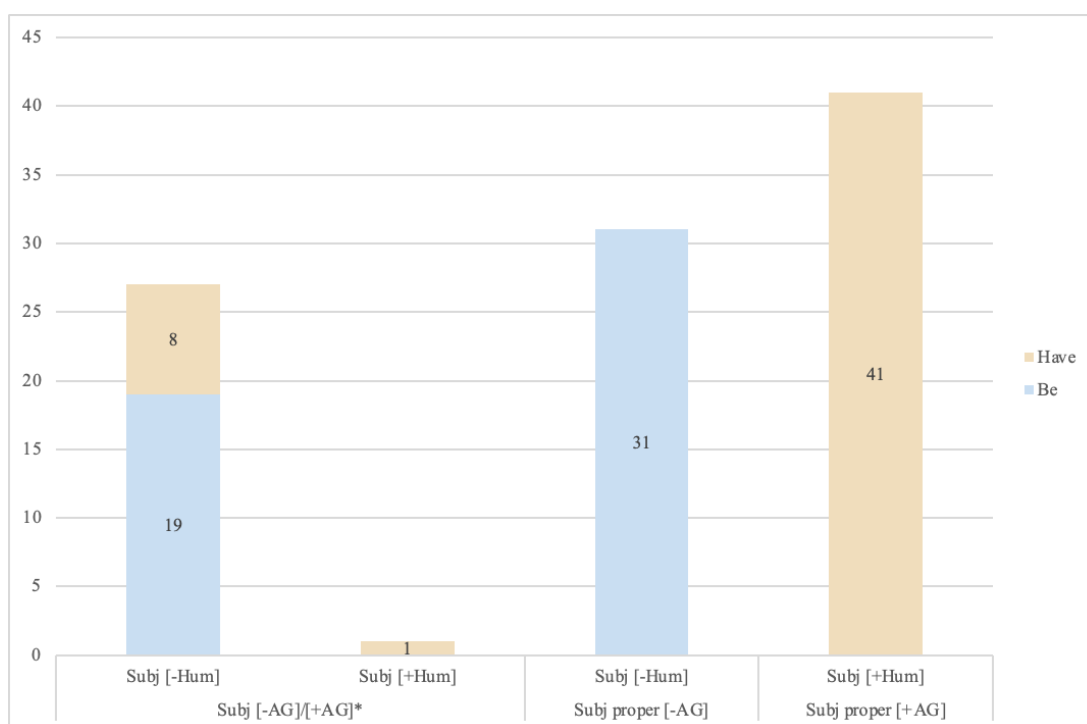


FIGURE 3.26 – Correlation of ‘have’ and ‘be’ with agentive/internal cause/non-agentive subjects and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-AG]/[+AG] *[-Hum]</b>	19	8	27
<b>Subj [-AG]/[+AG] *[+Hum]</b>	/	1	1
<b>Subj proper [-AG] [-Hum]</b>	31	/	31
<b>Subj proper [+AG] [+Hum]</b>	/	41	41
<b>Total</b>	49	51	100

TABLE 3.35 – Frequency of ‘have’ and ‘be’ based on agentive/internal cause/non-agentive subjects and human/non-human subjects.

Human subjects can be classified into two categories: genuinely agentive (103) or non-agentive yet causative (104), and in both scenarios, they opt for ‘have’:

- (103) [...] dopo aver letto l'articolo ho proceduto  
 after have.INF read.PTCP DEF.ART.article have.PRS.1SG proceed.PTCP  
 come indicato.  
 as indicate.PTCP

'[...] After reading the article I proceeded as indicated.'

- (104) [...] ero certo che a vela avrebbero proceduto più lentamente  
 be.IPFV.1SG sure that under.sail have.COND.3PL proceed.PTCP more slowly  
 di me.  
 than me

'[...] I was certain that under sail they would proceed more slowly than I did.'

In (103), the subject is fully agentive as it has complete control over the action, while in (104), the subject does not deliberately choose progress more slowly. This is similar to the expression *iniziare alla grande* 'to start big': the subject may wish to start ambitiously, but the adverbial conjunction reduces its agency.

When the subject is non-human, 'be' is the auxiliary used both when it is not agentive (105) and when it is internal cause (106). However, the internal cause factor allows for the use of 'have' with non-human subjects too, such as in (107), where the subject is an instrument:

- (105) [...] la discussione è proceduta come segue [...].  
 DEF.ART discussion be.PRS.3SG proceed.PTCP.F.SG as follow.PRS.3SG

'[...] the discussion proceeded as follows [...].'

- (106) Secondo me i lavori sono proceduti speditamente.  
 for me DEF.ART works be.PRS.3PL proceed.PTCP.M.PL smoothly

'In my opinion, so far the work has proceeded smoothly'.

- (107) La ruspa [...] ha proceduto fra le 20 e  
 DEF.ART excavator have.PRS.3SG proceed.PTCP between DEF.ART 8 p.m. and  
 le 21:30 alla demolizione [...] di Palazzo Montagna.  
 DEF.ART 9:30 p.m. to.DEF.ART demolition of P.M.

'The excavator [...] proceeded between 8 p.m. and 9:30 p.m. to the demolition, partial but final, of Palazzo Montagna.'

The example in (107) may appear contradictory at first glance, but the interpretation lies within the nuanced interplay between the lack of agentivity and the presence of causativity, provided

by the concept of internal cause. Moreover, the choice also hinges on the strength of the subject’s association with human traits: in (106), the subject *lavori* can be interpreted both as events occurring and as actions performed by someone represented by the deverbal noun.

However, this link is less robust than in (107), where the *ruspa* ‘excavator’ obviously cannot proceed without a person operating it. In this case, what can be termed as ‘shifted intentionality’ occurs, wherein the action of proceeding is attributed to the inanimate item but ultimately refers to the people controlling it. The human trait behind this intensifies the causativity and can lead us to interpret this sentence as a sort of indirect transitive construction.

**Adverbials** Among the most frequent adverbials are those indicating manner and adverbial arguments. However, these adverbials themselves do not influence the auxiliary distribution. In the case of adverbials indicating manner, the selection still depends on the human trait, while for adverbial arguments, they only intensify the causative role of the subject.

**ADVERBIAL OF MANNER** The adverbials of manner, such as *faticosamente* ‘laboriously’, are among the most frequent adverbials, totaling 59 occurrences. Human subjects always select ‘have’, while for non-human subjects, ‘be’ is the canonical auxiliary, as depicted in Figure 3.27.

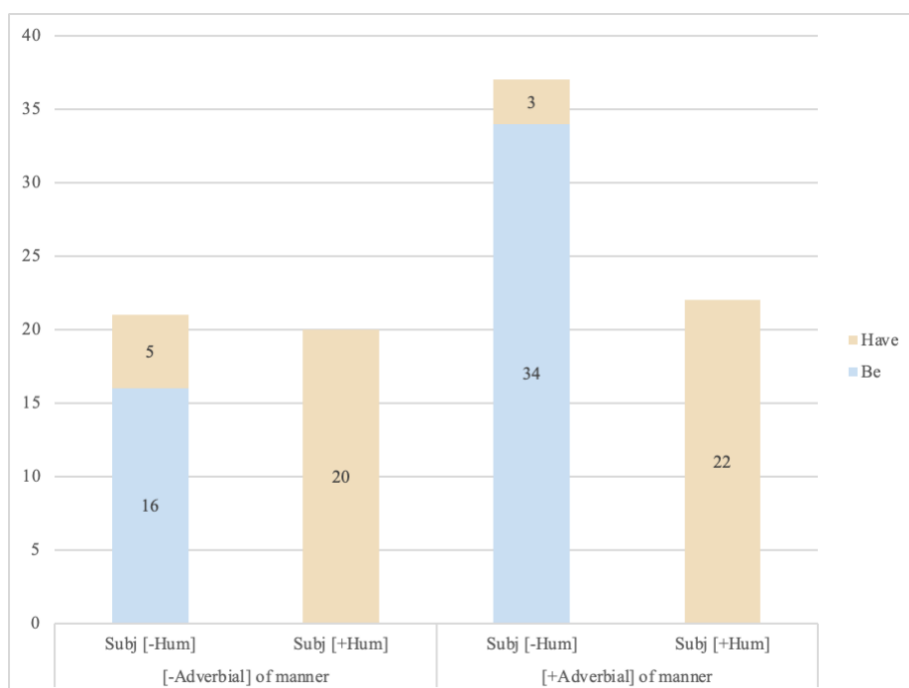


FIGURE 3.27 – Correlation of ‘have’ and ‘be’ with adverbial of manner and human/non-human subjects.

Factors	Data 'Be'	Data 'Have'	Total
[-Adverbial of manner] Subj [-Hum]	16	5	21
[-Adverbial of manner] Subj [+Hum]	/	20	20
[+Adverbial of manner] Subj [-Hum]	34	3	37
[+Adverbial of manner] Subj [+Hum]	/	22	22
<b>Total</b>	50	50	100

TABLE 3.36 – Frequency of ‘have’ and ‘be’ based on the adverbial of manner and human/non-human subjects.

Figure 3.27 and Table 3.36 illustrate how the human feature is decisive for the selection of the auxiliary. The presence or absence of adverbials of manner does not seem to play any role as such, unless it is preferentially associated with a human subject. Examples (108) and (109) exemplify canonical cases.

- (108) Anche durante la Meditazione camminata hanno proceduto  
 also during DEF.ART Meditation walking have.PRS.3PL proceed.PTCP  
 silenziosi, concentrati, camminando lentamente [...].  
 silent concentrated walking slowly

‘Also during the Walking Meditation they proceeded silently, concentrated, walking slowly [...].’

- (109) [...] se le cose fossero procedute normalmente, avrei  
 if DEF.ART things be.PST.3PL proceed.PTCP.F.PL normally have.COND.1SG  
 scelto una nascita in casa.  
 choose.PTCP INDF.ART birth a.home

‘[...] if things had proceeded normally, I would have chosen a home birth.’

In (108), the adverbial of manner is represented by *silenziosi* ‘silently’, an adjective indicating the manner in which the action was carried out. This is further emphasized by the adjective *concentrati* and the gerund *camminando lentamente*, which both describe the manner of walking.

**ADVERBIAL ARGUMENTS** The adverbial argument is represented by the complement introduced by *a* in the expression *procedere a* ‘proceed to’. This adverbial appears 15 times and is associated with causativity: when the subject is human, it is also agentive, and when the subject is non-human, it denotes internal cause. In both cases, as seen in (110) and (111) respectively, the auxiliary used with the adverbial argument is exclusively ‘have’.

The choice of ‘have’ can be further explained if we consider the hypothesis that this construction is an indirect transitive one.

(110) Di recente il sottoscritto amministratore ha proceduto  
 recently DEF.ART undersigned administrator have.PRS.3SG proceed.PTCP

all’abbattimento degli alberi [...].  
 at.DEF.ART cutting of.DEF.ART trees

‘Recently, the undersigned administrator proceeded to cut down trees [...].’

(111) Gli aerei hanno proceduto all’avvistamento  
 DEF.ART airplanes have.PRS.3PL proceed.PTCP at.DEF.ART.sighting

dell’imbarcazione [...]  
 of.DEF.ART.vessel

‘The airplanes proceeded to sight the vessel [...].’

The sentences in (110) and (111) are similar, but the distinction between agentive and internal cause factors depends on the human trait of the subject. The expression *procedere a* requires causativity, which can manifest as agentivity in the case of a human subject (110), whereas it is solely causativity in the case of a non-human subject (111). In (111), *aerei* (airplanes) acts as an instrument with shifted intentionality, meaning that the causative action originates from a human entity but is applied to the inanimate subject.

To conclude, the dynamic verb *procedere* is influenced by the human trait and the resulting agentivity or lack thereof. Agentivity is associated with human subjects, which take ‘have’ as their auxiliary, while non-human subjects with *procedere* are predominantly non-agentive and select ‘be’.

Additionally, the internal cause factor may lead to either ‘have’ or ‘be’ depending on the subject’s nature: ‘have’ is preferred with human subjects, while ‘be’ is favored with non-human subjects, although ‘have’ remains possible when causativity is emphasized. Other factors, such as adverbials, do not have a decisive impact unless associated with the human or non-human nature and causativity.

Table 3.37 outlines the primary factors influencing the auxiliary selection for *procedere*, which align with the patterns observed for *iniziare* and *cominciare*.

Agentivity	Human	Auxiliary
Internal Cause/Agentive	Human →	Have
Internal Cause/Non-Agentive	Non-human →	Be

TABLE 3.37 – Primary factors influencing the distribution of auxiliaries for *procedere*.

### 3.7.4 VERBS INDICATING THE CONTINUATION OF AN ACTION: *continuare*

*Continuare* ‘go on’, ‘carry on’ exhibits a comparable frequency of occurrences in the ItTenTen (16) corpus: 49% of instances with ‘have’ and 51% with ‘be’ when filtering out nouns using part-of-speech filtering, with 377 occurrences of ‘have’ and 399 occurrences of ‘be’. As observed with other dynamic verbs, the human trait exerts a significant influence.

**Human and non-human subjects** Auxiliary distribution is significantly influenced by the human trait: non-human subjects tend to prefer ‘be’ while human subjects favor ‘have’, as depicted in Figure 3.28 and Table 3.38, based on 95 occurrences of the intransitive construction (without direct object).

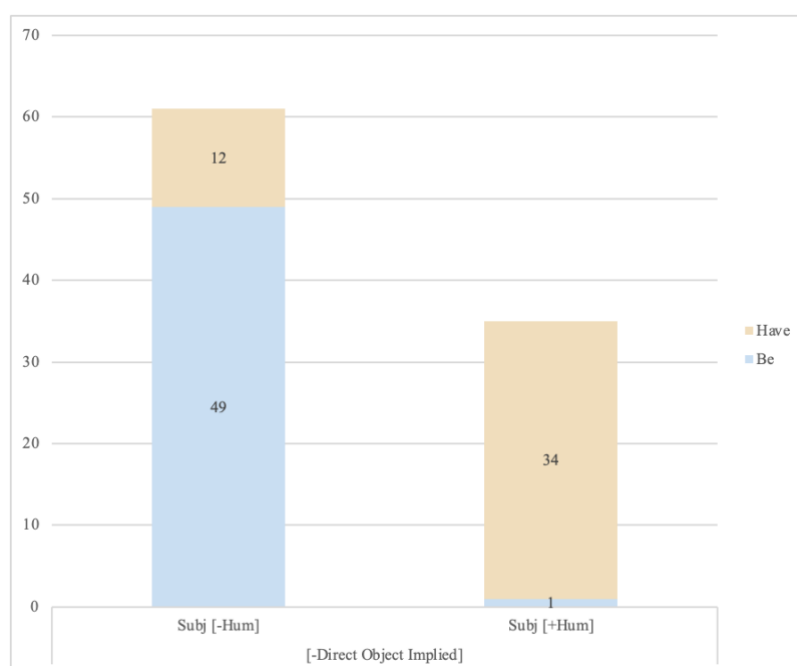


FIGURE 3.27 – Correlation of ‘have’ and ‘be’ with human/non-human subjects in the absence of direct object.

Factors	Data ‘Be’	Data ‘Have’	Total
[-DO] Subj [-Hum]	49	12	61
[-DO] Subj [+Hum]	1	33	34
<b>Total</b>	50	45	95

TABLE 3.38 – Frequency of ‘have’ and ‘be’ based on human/non-human subjects in the absence of direct object.

Examples with a non-human subject and ‘be’ and with human subject with ‘have’ are provided in respectively (112) and (113):

(112) [...] la lezione è continuata come se nulla fosse  
DEF.ART class be.PRS.3SG continue.PTCP.F.SG as if nothing be.SUBJ.3SG

successo.  
happen.PTCP

‘The class continued as if nothing had happened.’

(113) I due attori hanno continuato parlando della chimica  
DEF.ART two actors have.PRS.3PL continue.PTCP talking of DEF.ART chemistry

instauratasi durante le riprese [...].  
establish.PTCP.F.SG-REFL during DEF.ART filming

‘The two actors went on to talk about the chemistry established during the filming [...].’

The sentences in (112) and (113) illustrate different uses of *continuare*. In (112), the structure represents a more typical use of *continuare*, while (113) features an intransitive use with a human, agentive subject. In (113), *continuare* functions as a sort of semi-auxiliary, with the gerund *parlando* ‘talking’ specifying how the subject continued the action.

Agentivity, internal cause, and lack of agentivity are other factors correlated to the feature  $\pm$  HUMAN conveyed by the subject.

**Agentivity, internal cause and lack of agentivity** When considering the factor of agentivity, human subjects are predominantly agentive and co-occur with ‘have’, while non-agentive subjects are typically non-human and tend to select ‘be’. However, among the non-human subjects, there are 19 occurrences with an internal cause, leading to an equal frequency of ‘have’ and ‘be’. This tendency is depicted in Figure 3.28 and Table 3.39.

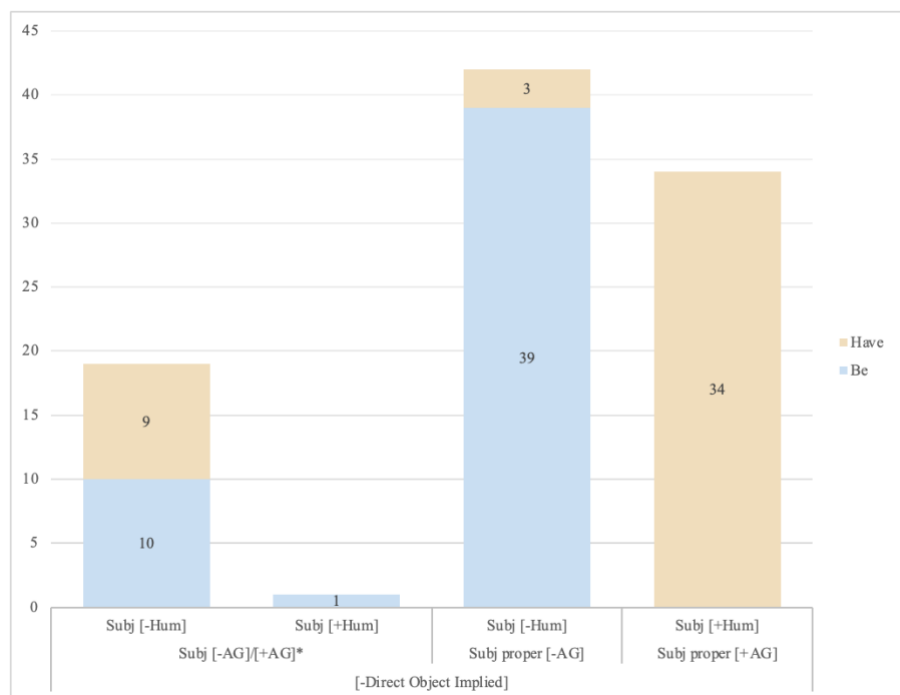


FIGURE 3.28 – Correlation of ‘have’ and ‘be’ with agentive/non-agentive and internal cause subjects, and human/non-human subjects

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] Subj [-AG]/[+AG] * [-Hum]</b>	10	9	19
<b>[-DO] Subj [-AG]/[+AG] * [+Hum]</b>	1	/	1
<b>[-DO] Subj proper [-AG] [- Hum]</b>	39	3	42
<b>[-DO] Subj proper [+AG] [+ Hum]</b>	/	34	34
<b>Total</b>	50	46	96

TABLE 3.39 – Frequency of ‘have’ and ‘be’ according to agentive/non-agentive and internal cause subjects, and human/non-human subjects.

The tendency of auxiliary selection of *continuare* aligns with the other verbs referring to a stage of process analyzed so far: ‘be’ with non-human non-agentive subjects, e.g. (114), and ‘have’ with a human agentive subject, e.g. (115).

However, as already observed in the case of *procedere*, with *continuare* the use of both ‘be’ (116) and ‘have’ (117) is possible with an internal cause non-human subject, and even to a larger extent.

(114) La serata è continuata piacevolmente divertendoci [...].  
 DEF.ART evening be.PRS.3SG continue.PTCP.F.SG pleasantly having.fun.REFL

‘The evening continued pleasantly as we had fun [...].’

(115) L'assessore ha continuato affermando che quanto  
 DEF.ART .councilor have.PRS.3SG continue.PTCP saying that what

fatto fino ad ora è solo un punto di partenza.  
 do.PTCP so.far be.PRS.3SG only INDF.ART point of departure

'The councilor went on to say that what has been done so far is only a starting point.'

(116) Dopo aver scoperto che la foto di gruppo [...] non  
 after have.INF discover.PTCP that DEF.ART photo of group [...] not

comprendeva il poeta di Marradi [...] le ricerche sono  
 include.IPFV.3SG DEF.ART poet of Marradi [...] DEF.ART search be.PRS.3PL

continue, fino a scoprire una nuova foto.  
 continue.PTCP.F.PL until find.INF INDF.ART new photo

'After discovering that the group photo [...] did not include the poet from Marradi [...] the search continued until a new photo was found.'

(117) Numerosi studi hanno continuato nel 1990 e anche  
 numerous studies have.PRS.3PL continue.PTCP in. DEF.ART 1990 and even

dopo la morte del Dottor Atkins.  
 after DEF.ART death of. DEF.ART Dr. Atkins

'Numerous studies continued in 1990 and even after Dr. Atkins' death.'

(116) and (117) are particularly interesting cases as they involve non-human subjects with internal cause, characterized by shifted intentionality. These examples feature inanimate subjects with a clear connection to humans: the research or studies continue because of the people conducting it. The continuation is further emphasized in (117) by the adjunct *e anche dopo la morte del Dottor Atkins* 'and even after Dr. Atkins' death'.

A similar scenario arises with subjects such as vehicles (118), which cannot be considered animate but are instruments with causative characteristics attributed to them by the person driving them. This makes them inanimate yet characterized by internal cause due to shifted intentionality.

(118) Dopo una breve sosta [...], il veicolo aveva continuato su...  
 after INDF.ART brief stop [...] DEF.ART vehicle have.IPFV.3SG continue.PTCP on

ma solo per un minuto.  
 but only for INDF.ART minute

'After a brief stop [...], the vehicle had continued... but only for a minute.'

As a result, in situations where agentivity is absent but causativity is present, there tends to be more variation in the choice of auxiliary. When internal causation is involved and/or there is some perceived connection to animacy (as seen in cases of shifted intentionality and examples involving instruments such as vehicles), the likelihood of selecting ‘have’ increases.

Moreover, considering that ‘have’ is particularly correlated with agentivity and causativity, there is a certain circular relationship between auxiliaries and causativity: causativity explains the presence of ‘have’, conversely, the presence of ‘have’ induces the perception of a non-human subject as being internally causal when certain elements permit it (such as the semantics of the verb or the possible link of the subject to a human referent).

**Adverbials** The most frequent adverbials are those of manner (e.g., *con costanza* ‘with constancy’) and time (e.g., *nel 2016* ‘in 2016’).

**ADVERBIALS OF MANNER** The presence of adverbials of manner or time does not influence the choice between ‘have’ and ‘be’. Instead, it is the human nature of the subject and the associated agentivity or lack thereof that determine the auxiliary selection. Non-human subjects accompanied by adverbials of manner are typically non-agentive and select ‘be’, whereas human subjects accompanied by such adverbials are usually agentive and select ‘have’. These correlations are represented in Figure 3.29 and exemplified in (119) and (120). They are in line with the general trends observed in the data without the presence of an adverb of manner (Figures 3.29 & 3.40).

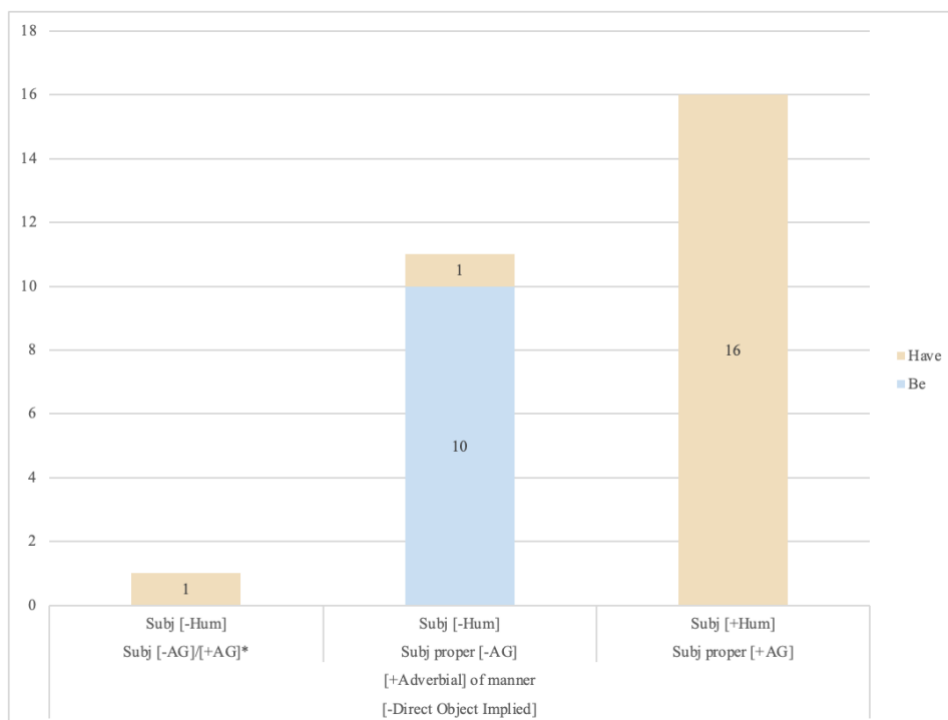


FIGURE 3.29 – Correlation of ‘have’ and ‘be’ with adverbial of manner, agentive/non-agentive and internal cause subjects, and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] [+Adverbial of manner] Subj [-AG]/[+AG] * [-Hum]</b>	/	1	1
<b>[-DO] [+Adverbial of manner] Subj proper [-AG] [-Hum]</b>	10	1	11
<b>[-DO] [+Adverbial of manner] Subj proper [+AG] [+Hum]</b>	/	16	16
<b>Total</b>	10	18	28

TABLE 3.40 – Frequency of ‘have’ and ‘be’ based on adverbial of manner, agentive/non-agentive and internal cause subjects, and human/non-human subjects.

(119) [...] la celebrazione è continuata come di consueto [...].  
 DEF.ART celebration be.PRS.3SG continue.PTCP.F.SG as usual

‘[...] The celebration continued as usual [...].’

(120) Con costanza ho continuato e ho cominciato a  
 with constancy have.PRS.1SG continue.PTCP and have.PRS.1SG start.PTCP to  
 vedere dei risultati.  
 see.INF of.DEF.ART results

‘Consistently, I continued and started to see some results.’

**TIME ADVERBIALS** The analysis conducted on adverbials of manner also extends to time adverbials. Out of the 21 occurrences, there is a distinct preference for ‘be’, especially with non-human and non-agentive subjects (8 occurrences with ‘be’ vs. 1 with ‘have’). However, more instances with ‘have’ are observed when internal cause is involved (4 ‘have’ vs. 1 ‘be’).

Conversely, when the subject is characterized by the human trait and agentivity, the use of ‘have’ dominates (all 7 occurrences with ‘have’). Thus, the primary influencing factors for auxiliary selection remain the human trait and the presence or absence of agentivity, demonstrating the nuanced interplay between these factors.

To summarize, the dynamic verb *continuare*, which signifies the continuation of an action, follows a pattern similar to other dynamic verbs in the sense that the auxiliary selection depends most of all on the features of agentivity which correlates often but not always with the feature ‘human’, as represented in Table 3.41.

Agentivity	Human	Auxiliary
Agentive	Human →	Have
Non- agentive	Non-human →	Be
Internal Cause	Non-human →	Have/Be

TABLE 3.41 – Primary factors influencing the distribution of auxiliaries for *continuare*.

### 3.7.5 VERBS INDICATING THE CONTINUATION OF A PROCESS: *prosequire*

*Prosequire* ‘go on’ is another verb used to express the continuation of an action. In ItTenTen (16), it exhibits a higher frequency with ‘be’ compared to ‘have’: 36% of instances with ‘have’ compared to 64% with ‘be’ when the part-of-speech filter is applied, with 297 occurrences of ‘be’ and 170 of ‘have’.

**Human and non-human subjects** As observed with other dynamic verbs in this study, *prosequire* is strongly influenced by the animacy of the subject: ‘have’ is used when the subject is human, whereas ‘be’ is preferred when the subject is non-human. This tendency is illustrated in Figure 3.30 and Table 3.42.

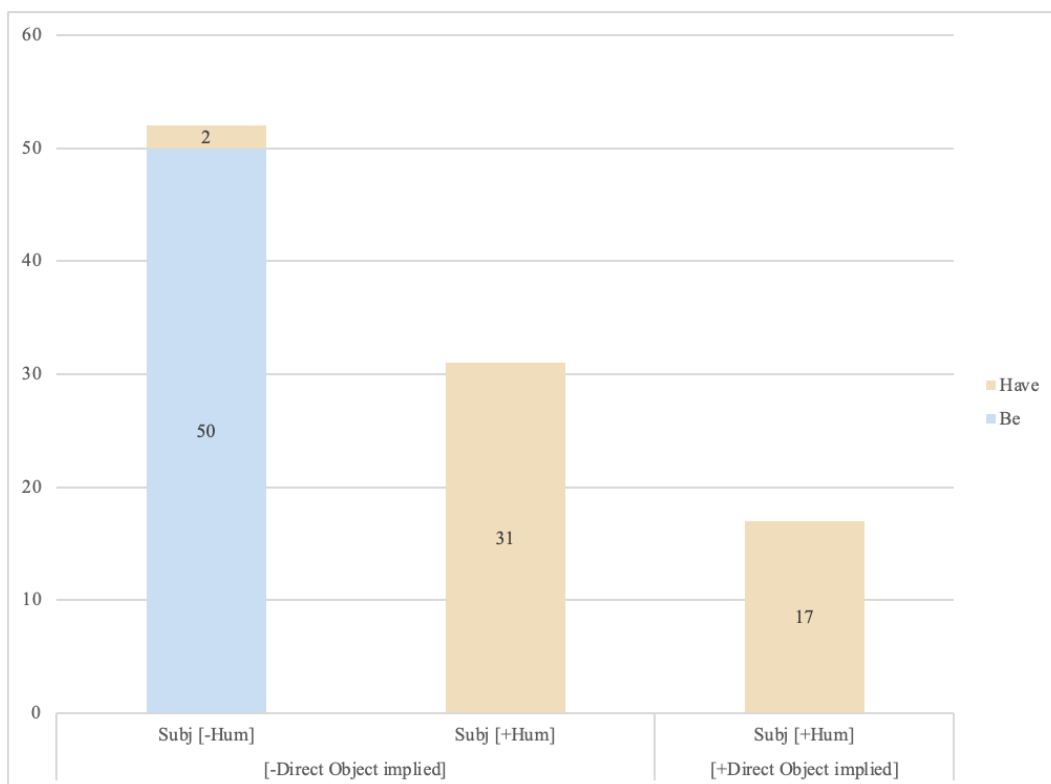


FIGURE 3.30 – Correlation of ‘have’ and ‘be’ with the presence/absence of direct object and a human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
[-DO] Subj [-Hum]	50	2	52
[-DO] Subj [+Hum]	/	31	31
[+DO] Subj [+Hum]	/	17	17
<b>Total</b>	50	50	100

TABLE 3.42 – Frequency of ‘have’ and ‘be’ based the presence/absence of direct object and human/non-human subjects.

As indicated in Table 3.42 and Figure 3.30, there are 17 instances where *proseguire* is accompanied by a implicit direct object. In these contexts, the verb takes on the meaning of ‘say’ or ‘continue speaking’. Consider (121) as an example:

(121) Rivolto ad alcuni esponenti politici, l’Assessore ha  
 address.PTCP to some figures politicals DEF.ART.Assessor have.PRS.3SG

proseguito: “Dicano [...]”.  
 proceed.PTCP say.SUBJ.3PL

‘Addressing some political figures, the Assessor proceeded: “They say...[...]”.’

In this scenario, the auxiliary is clearly ‘have’. Another context where ‘have’ is used as the auxiliary is with human subjects in the intransitive context, as seen in (122) and (123):

(122) Ma la scrittrice ha proseguito precisando: “Per questo [...]”.  
 But DEF.ART writer have.PRS.3SG go.on.PTCP clarifying for this

‘But the writer went on to clarify, “For this [...]”.’

(123) Alla fine ho proseguito, ma solo fino a Tardajos [...].  
 in.DEF.ART end have.PRS.1SG go.on.PTCP but only until Tardajos

‘In the end I continued, but only as far as Tardajos [...]’

While (122) presents a similar meaning to (121), namely ‘continue speaking’, it also implies ‘go on’ as emphasized by the gerund indicating the manner in which the subject went on. The verb *proseguire*, therefore, behaves like a sort of semi-auxiliary. In contrast, (123) implies going on in space.

When the subject is non-human, however, the canonical auxiliary is ‘be’, as seen in (124):

(124) Ci siamo conosciuti in treno [...] e il nostro viaggio  
 REFL be.PRS.1PL know.PTCP.M.PL on train and DEF.ART our journey

è proseguito insieme!  
 be.PRS.3SG continue.PTCP together

‘We met on the train [...] and our journey continued together!’

The only two occurrences with a non-human subject that present ‘have’ are characterized by agentivity or internal cause due to shifted intentionality. Consequently, the examples presented thus far also reflect the concept of agentivity: when the subject is human, it is typically agentive, whereas when it is non-human, it is predominantly non-agentive.

***Agentivity, lack of agentivity and internal cause*** As the human trait reflects agentivity, this factor becomes consequently important as well. Agentivity correlates with ‘have’, while the lack of it correlates with ‘be’, as depicted in Figure 3.31 and Table 3.43.

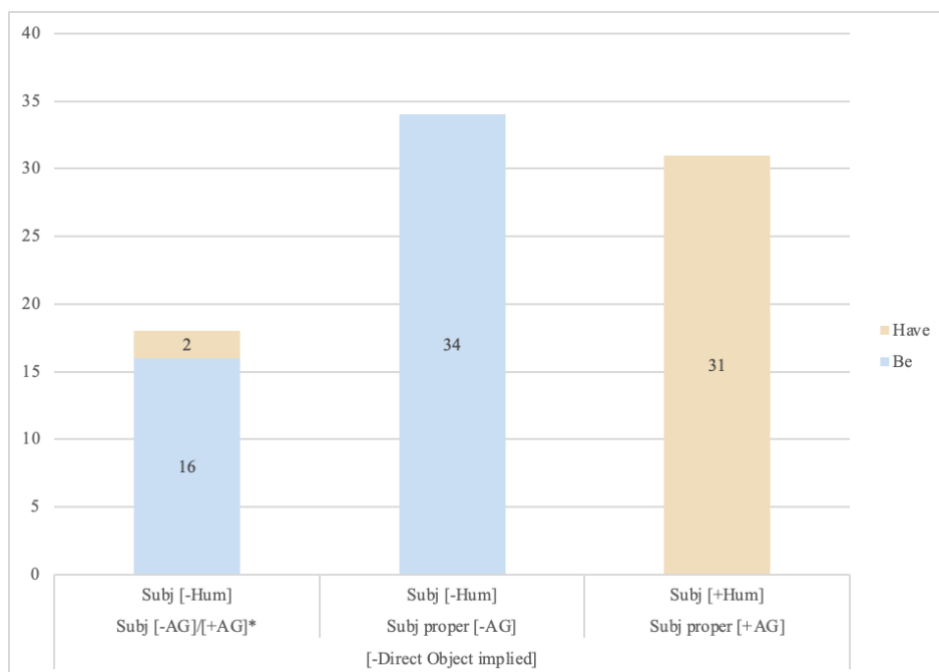


FIGURE 3.31 – Correlation of ‘have’ and ‘be’ with agentive/non-agentive/ internal cause subjects and human/non-human subjects in the absence of direct object.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-DO] [-AG]/[+AG] *[-Hum]</b>	16	2	18
<b>[-DO] Subj proper [-AG] [-Hum]</b>	34	/	34
<b>[-DO] Subj proper [+AG] [+Hum]</b>	/	31	31
<b>Total</b>	50	33	83

TABLE 3.43 – Frequency of ‘have’ and ‘be’ based on agentive/non-agentive/ internal cause subjects and human/non-human subjects in the absence of direct object.

Therefore, human agentive subjects select ‘have’ (125), while non-human non-agentive subjects select ‘be’ (126).

(125) Abbiamo proseguito per la costiera, fino ad arrivare  
 have.PRS.1PL proceed.PTCP along DEF.ART coast until come.PTCP

alla bellissima città di Brighton.  
 at. DEF.ART beautiful city of B.

‘We proceeded along the coast until we reached the beautiful city of Brighton.’

(126) La piacevole serata è proseguita con la “Sonata  
 DEF.ART enjoyable evening be.PRS.3SG proceed.PTCP.F.SG with DEF.ART Sonata

in mi bemolle KV481” di Wolfgang Amadeus Mozart.  
 in mi bemolle KV481 of W. A. Mozart

‘The enjoyable evening continued with Wolfgang Amadeus Mozart’s “Sonata in mi bemolle maggiore KV481”.’

Thus, human trait is the primary criterion that determines auxiliary selection. Secondly, there is agentivity: full agentivity requires ‘have’, whereas non-agentive subjects go along with the auxiliary ‘be’. Internal cause subjects prefer ‘be’, but can accommodate both auxiliaries due to the intermediate interpretation between lack of agentivity and causative behavior of the subject. Consider (127) as an example with ‘have’ and (128) as an example with ‘be’:

(127) Se la corriera avesse proseguito, che cosa sarebbe  
 if DEF.ART bus have.SUBJ.3SG proceed.PTCP what be.COND.3SG  
 avvenuto?  
 happen.PTCP

‘If the bus had proceeded, what would have happened?’

(128) Così gli studi sono proseguiti, e si arrivò a  
 So DEF.ART studies be.PRS.3SG proceed.PTCP.M.PL and REFL come.PST.3SG to  
 stabilire che anche l’udito era colpito.  
 determine.PTCP that even DEF.ART.hearing be.IPFV.3SG affect.PTCP

‘So the studies continued, and it came to be determined that hearing was also affected.’

It can be assumed that (127) has a strong correlation with ‘have’ because *corriera* ‘bus’ is an instrument presenting internal cause due to shifted intentionality from the person controlling it.

However, in (128), the subject is causative, similar to those subjects that appeared with ‘have’ in *continuare*, but ‘be’ is used, since the subject is also affected. It may be due to the semantics of *proseguire* itself: it has been seen that when ‘have’ is used, as in (127) or with human subjects (125), the meaning is of ‘continue in space’ (as well as that of ‘speak’, this meaning being compatible only with human subjects).

**Adverbials** The adverbials do not influence the auxiliary selection with *proseguire*. Although the most frequent adverbials include adverbials of manner, adverbial arguments, and time adverbials with or without duration expressed, the distribution of auxiliaries still depends on the human trait or agentivity of the subject: non-human, non-agentive subjects correlate with ‘be’, while human, agentive subjects correlate with ‘have’.

To conclude, the verb *proseguire* is in line with the patterns observed with other dynamic verbs analyzed thus far, where the human trait and agentivity emerge as the essential factors influencing the distribution of ‘have’ and ‘be’, as illustrated in Table 3.44.

Agentivity	Human	Auxiliary
Agentive	Human →	Have
Non-agentive	Non-human →	Be

TABLE 3.44 – Primary factors influencing the distribution of auxiliaries for *proseguire*.

Note that in the verbs analyzed so far, variation of ‘have’ and ‘be’ arises when the subject has both the non-human and the internal-cause feature.

### 3.7.6 VERBS INDICATING THE COMPLETION OF A PROCESS: *finire*

The verb indicating the completion of an action, *finire*, tends to have significantly more instances with ‘be’ than with ‘have’. In the ItTenTen (16) corpus, there are 4% instances of ‘have’ (170 occurrences) compared to 96% of ‘be’ (10,003 occurrences) when the transitive context is excluded as much as possible using the part-of-speech filter.

However, in the 100-sentences corpus, the distribution of ‘have’ or ‘be’ remains dependent on the human trait.

***Human and non-human subjects*** Human subjects correlate with ‘have’, and non-human subjects with ‘be’, as depicted in Figure 3.32 and Table 3.45

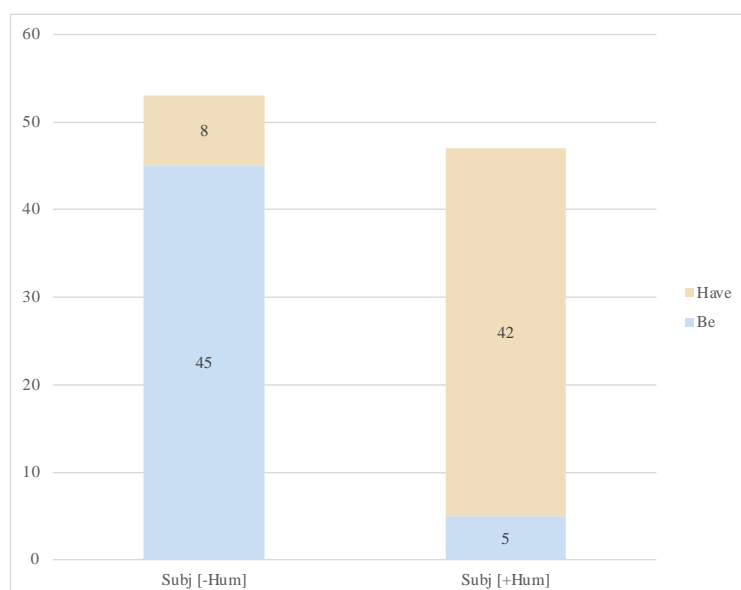


FIGURE 3.32 – Correlation of ‘have’ and ‘be’ with human/non-human subjects.

Factors	Data 'Be'	Data 'Have'	Total
Subj [-Hum]	45	8	53
Subj [+Hum]	5	42	47
Total	50	50	100

TABLE 3.45 – Frequency of ‘have’ and ‘be’ based human/non-human subjects.

As shown in Figure 3.32 and Table 3.45, the auxiliary selection does not depend on the presence or absence of an implied direct object, as all the occurrences were intransitive. The canonical auxiliary with non-human subjects is ‘be’ (129), while the canonical auxiliary with human subjects is ‘have’ (130):

(129) La mia missione è finita.  
DEF.ART my mission be.PRS.3SG finish.PTCP.F.SG

‘My mission is done.’

(130) Qui per fortuna ho finito, ora ce la squagliamo!  
here luckily have.PRS.1SG finish.PTCP now LOC it scam.PRS.1PL

‘Here, luckily, I’m done, now let’s scam!’

The intransitive construction occurs more often with a non-human subject, as in (129), and tends to select in this case ‘be’. Conversely, when it combines with a human subject, the auxiliary ‘have’ is mostly selected. However, (130) illustrates that even with human subjects, *finire* can occur in an intransitive construction. The opposition between human vs non-human subjects is evidently correlated with the parameter of agency.

**Agentivity, lack of agentivity and internal cause** Agentivity emerges as the second significant factor influencing the distribution of auxiliaries. It exhibits a strong correlation with ‘have’, while lack of agentivity is associated with ‘be’, observed in both human and non-human subjects.

The typical scenario involves non-human, non-agentive subjects, making the human trait the primary determinant. However, in the presence of internal cause, while human subjects predominantly favor ‘have’, non-human subjects display variation, as illustrated in Figure 3.33.

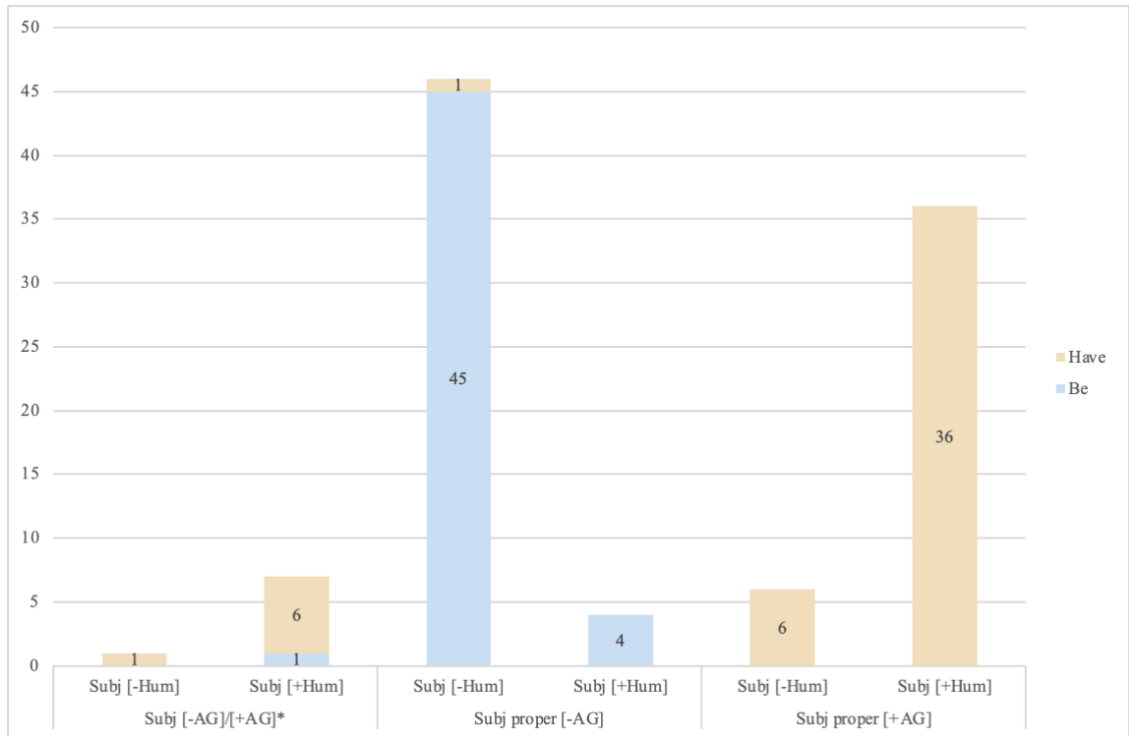


FIGURE 3.33 – Correlation of ‘have’ and ‘be’ with agentive/non-agentive/ internal cause subjects and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-AG]/[+AG] *[-Hum]</b>	/	1	1
<b>Subj [-AG]/[+AG] * [+Hum]</b>	1	6	7
<b>Subj proper [-AG] [-Hum]</b>	45	1	46
<b>Subj proper [-AG] [+Hum]</b>	4	/	4
<b>Subj proper [+AG] [-Hum]</b>	/	6	6
<b>Subj proper [+AG] [+Hum]</b>	/	36	36
<b>Total</b>	50	50	100

TABLE 3.46 – Frequency of ‘have’ and ‘be’ based on agentive/non-agentive/ internal cause subjects and human/non-human subjects.

As both the Figure 3.33 and Table 3.46 show, the most frequent scenarios are the non-human non-agentive subjects selecting ‘be’ (131) and human agentive subjects selecting ‘have’ (132):

(131) [...] il sondaggio è finito il: 05 Ott 2014.  
 DEF.ART survey be.PRS.3SG finish.PTCP DEF.ART 05 Oct 2014

‘[...] The survey ended on: 05 Oct 2014.’

- (132) Dopo che i soldati hanno finito e se ne sono  
 after that DEF.ART soldiers have.PRS.3PL finish.PTCP and REFL of.it be.PRS.3PL  
 andati i PG possono continuare il viaggio [...].  
 go.PTCP.M.PL DEF.ART PG can.PRS.3PL continue.PTCP DEF.ART journey  
 ‘After the soldiers have finished and left the PGs can continue the journey [...].’

When the subject lacks proper agency, the auxiliary ‘be’ is consistently chosen, even with human subjects. There is a noteworthy case in this regard (133):

- (133) [...] quando sono tornata a casa ero finita!!!  
 when be.PRS.1SG come.back.PTCP.F.SG at home be.IPFV.1SG finish.PTCP.F.SG  
 ‘[...] when I got home, I was finished!!!!’

(133) is the quintessential example one thinks of while discussing the verb *finire*, reflecting the semantics of ‘finish’ when the subject is human. In this case, the verb has a complex relationship with the adjective *sfinito* ‘exhausted’.

The role of agentivity or internal cause, particularly when combined with the human or animate trait, is especially pronounced. ‘Have’ is preferred when the subject is human and presents internal cause (134), and even when the subject is non-human and agentive due to the verb’s semantics (as in 135, where the subject is an instrument).

When the subject is an internal cause due to shifted intentionality (136), ‘have’ also appears, even if there is only one occurrence.

- (134) [...] sono nato per ultimo, tutti pensavano che mamma  
 be.PRS.1SG born.PTCP for.last everyone think.IPFV.1PL that mom  
 avesse finito, invece eccomi qua.  
 have.SUBJ.3SG finish.PTCP but here.I.am  
 ‘[...] I was born last, everyone thought mom was done, but here I am.’

- (135) ora scappo... la lavatrice ha finito, devo  
 now run.PRS.1SG DEF.ART washing.machine have.PRS.3SG finish.PTCP have.to.PRS.1SG  
 stendere il bucato [...].  
 hang.out.INF DEF.ART laundry  
 ‘I run now...the washing machine is finished; I have to hang out the laundry [...].’

(136) [...] sistemate le guarnizioni del battiporta, il lavoro  
 fix.PTCP.F.PL DEF.ART gaskets of. DEF.ART door.knocker DEF.ART job  
 è finito.  
 be.PRS.3SG finish.PTCP

‘[...] fixed the door knocker gaskets, the job is done.’

In (134), we observe how the subject, ‘mom’, is animate though not human. Because it lacks the ability to control the moment it gives birth, the subject is not properly agentive but still causative. In (135), the washing machine, a non-human subject, engages in a process where it cannot be the agent but must be the cause. Finally, in (136), the subject *lavoro* ‘job’) is still considered causative, as it is easily associated with human activity to the point where it stands for ‘be finished’.

As a result, the human/non-human distinction (or animate/non-animate) is the primary factor influencing auxiliary selection. Following this, causativity emerges as the sole determining factor within both categories. Notably, the more human the subject, the higher the tendency toward ‘have’ (it is important to note that causativity, in Italian, refers to the active role a subject can play in the action, and is therefore more closely related to ‘have’ and human subjects).

**Adverbials** The most frequent adverbials observed are adverbial arguments and time adverbials. However, their presence does not influence auxiliary selection. Instead, the choice of auxiliaries consistently depends on the semantic factors associated with the subject: the human trait and agentivity, as we have seen in the previous sections dedicated to other verbs referring to a stage of a process.

In conclusion, this dynamic verb, like others analyzed previously, exhibits a primary reliance on the human or non-human nature of the subject, which correlates with agentivity and lack of agentivity, respectively. Additionally, variation between ‘have’ and ‘be’ is possible when the subject is non-human and internal cause, as illustrated in Table 3.47

Agentivity	Human	Auxiliary
Agentive	Human →	Have
Non-agentive	Non-human →	Be
Internal Cause	Non-human →	Have/Be

TABLE 3.47 – Primary factors influencing the distribution of auxiliaries for *finire*.

Table 3.47 exhibits similarities to those presented for other verbs indicating a stage of a process. In some cases, the presence of internal cause allows for the selection of both auxiliaries, while for others, data are insufficient to identify a trend.

However, it remains consistent across all verbs that the human trait and agentivity are the primary factors influencing the choice between ‘have’ and ‘be’. Furthermore, variation between the two auxiliaries becomes possible when the subject’s internal cause nature facilitates an intermediate reading between lack of agentivity and causativity.

### 3.8 INTERMEDIARY RESULTS

The analysis has highlighted that the primary determinants for Italian peripheral verbs are the human nature of the subject and the agentivity factor. These factors, in conjunction with other relevant considerations such as the inherent semantics and aspect of the verb, determine the choice between ‘have’ and ‘be’ auxiliaries.

Initially, the verbs of the data sample were categorized based on their semantics, leading to classification according to their selection of ‘have’ or ‘be’ auxiliaries. The first group comprises verbs that primarily select ‘be’ as the auxiliary, although ‘have’ also occurs. The presence of the internal cause factor may influence auxiliary selection to varying degrees depending on the verb. These verbs are the change-of-state verb *cambiare*, as well as the verbs referring to various stages of a process *iniziare*, *cominciare*, *procedere*, *continuare*, *proseguire*, and *finire*.

The second group consists of verbs that predominantly select ‘have’ as the auxiliary. These verbs are characterized by semantics that conceive the subject as agentive or internal cause. While ‘have’ is the preferred auxiliary, ‘be’ may also be permitted, particularly in cases where there is a co-occurrence of non-agentive factors or the subject’s non-human nature. The verbs of this group are *cedere*, *contare*, *pesare*, *prevalere*, and *fallire*.

Finally, a third group is represented by the verb *suonare*, where the primary determinants are the internal cause nature of the subject and the aspect of the verb. In instances where the subject lacks agentivity, ‘be’ is consistently selected as the auxiliary. However, when the

subject acts as an internal cause, the aspect of the verb may influence the choice between ‘have’ (process) and ‘be’ (resultative).

Table 3.48 summarizes these predominant factors.

<b>Factors influencing auxiliary selection in Italian</b>		
<b>Aspects that could influence auxiliary distribution (the distribution of ‘have’ or ‘be’ is often the result of co-occurrence of more factors)</b>		
<b>1: HUMAN TRAIT</b>	<i>Animacy and human trait</i>	[+Hum] for ‘be’ selection verbs: ‘have’ [-Hum] for ‘be’ selection verbs: ‘be’
		[+Hum] for ‘have’ selection verbs: ‘have’ [-Hum] for ‘have’ selection verbs: ‘have’/’be’ (based on internal cause)
		[-Hum] for <i>suonare</i> : ‘have’/’be’ (based on internal cause and aspect)
<b>2: AGENTIVITY</b>	<i>Agentivity, lack of agentivity and internal cause behaviour</i>	[+AG] for ‘be’ selection verbs: ‘have’ [-AG] for ‘be’ selection verbs: ‘be’ [+Internal Cause] for ‘be’ selection verbs: ‘have’/’be’
		[+AG] for ‘have’ selection verbs: ‘have’ [-AG] for ‘have- selection verbs: ‘be’ [+Internal Cause] for ‘have’ selection verbs: ‘have’ (‘be’)
		[-AG] for <i>suonare</i> : ‘be’ [+Internal Cause] for <i>suonare</i> : ‘have’/ ‘be’
<b>3: ASPECT</b>	<i>Resultativity vs. Ongoing process</i>	<i>Resultativity for suonare</i> : ‘be’ <i>Process for suonare</i> : ‘have’/’be’

TABLE 3.48 – List of features that could affect the auxiliary selection

First, human trait (1) significantly affects the selection of ‘have’ or ‘be’ across all three groups. The first group of verbs (highlighted in yellow) exhibits a clear correspondence between human nature and the preference for ‘have’. In the other two groups, the selection of the auxiliary may also depend on the subject’s nature in terms of agentivity.

Second, agentivity (2) plays a crucial role, serving as the second key determinant. In the first group of verbs, there is a biunivocal correspondence between agentivity and ‘have’, as well as between the human nature of the subject and agentivity. In the other two groups, the

distribution depends on the interaction of the human trait and agentivity: if the subject is non-agentive, ‘be’ is selected as the auxiliary, whereas if it is human and internal cause, ‘have’ is preferred. If the subject is non-human, the internal cause factor allows for the possibility of ‘be’ as well.

Lastly, aspect (3) only pertains to the third group represented by the verb *suonare*. In this case, internal cause is the primary determinant for the potential usage of both ‘have’ and ‘be’. However, aspect may also play a role: when the construction implies an ongoing process interpretation, ‘have’ is favored, whereas ‘be’ may be preferred in other cases.

In conclusion, multiple factors interact in auxiliary selection, with internal cause, encompassing both non-agentivity and conveying causativity, being a key factor in the selection between ‘have’ and ‘be’. Furthermore, the choice of the auxiliary itself, such as ‘have’, may induce a causative reading, particularly when the subject is human. Verb’s semantics plays a significant role in this analysis, although syntax remains important. In the realm of peripheral verbs, after the initial distinction between transitive and intransitive, semantics often becomes more influential. Both semantics and syntax are essential, and they cannot be analyzed independently of each other.

As Loporcaro (2003) suggests<sup>17</sup>, certain semantic nuances may be challenging to analyze purely in syntactic terms, necessitating a consideration of both syntax and semantics. This is particularly evident in the case of peripheral verbs, where syntax alone may not fully account for the varied usage patterns observed.

Therefore, while syntax provides valuable insights semantics proves to be crucial for accounting auxiliary selection of peripheral verbs. By considering both syntax and semantics, a more comprehensive understanding of these verbs and their nuanced meanings can be achieved.

### 3.9 ASPECTUAL VERBS WITH INFINITIVE

Certain verbs referring to a stage of a process, specifically *iniziare*, *cominciare*, *continuare* and *finire* belong to a subset of verbs that, in addition to their role as main verbs, can also function as semi-auxiliaries when followed by a non-finite verb form possibly headed by a

---

<sup>17</sup> Loporcaro (2003: 213) discusses participial absolutes in Italian, presenting Dini’s proposal, according to which the acceptability of some participles with some unergatives can be explained in terms of Aktionsart: “Dini (1994) proposes a semantic (*Aktionsart*) constraint [...] that prevents atelic predicates from occurring in these constructions”. In our case, it is not just about *Aktionsart*, but about the general idea that semantics can explain some nuances while keeping syntax in mind.

preposition. In this role, they highlight a particular stage of the process expressed by the non-finite verb form and convey an aspectual meaning. An example is provided in (137).

- (137) Questa percentuale ha continuato a crescere.  
 This percentage have.PRS.3SG continue.PTCP to grow.INF

‘This percentage has continued to grow.’

Crucial in the context of this study is the hypothesis that the infinitive influences the choice between ‘have’ or ‘be’ of the semi-auxiliary of aspect.

To the best of my knowledge, this phenomenon has not been explicitly emphasized in works dedicated to auxiliary selection but has been identified under the label “restructuring” among one of the syntactic phenomena occurring in infinitive constructions with modal verbs, aspectual verbs, and motion verbs (Cinque 2004; Rizzi 1982; Rizzi 1976a). Rizzi (1976: 2) defines restructuring as an optional transformation of a verbal construction composed of two verbal predicates, namely an inflected verb and an infinitive, into a single verbal predicate.

This is how he explains that clitics corresponding to the arguments of the infinitive can also be positioned before the conjugated verb, and that the auxiliary selected by the infinitive verb can be transferred to the inflected verb (138 and 139).

- (138) Piero ha/è potuto venire con noi.  
 P. have/be.PRS.3SG can.PTCP come.INF with us

‘Piero could come with us’

- (139) La pioggia ha/è continuato/continuata ad aumentare [...].  
 DEF.ART rain have/be.PRS.3SG continue. PTCP/ PTCP.F.SG to increase.INF

‘The rain continued to increase [...].’

In this study, the verbs expressing a stage of a process, when combined with an infinitive, will rather be analyzed as semi-auxiliaries. The term semi-auxiliary is used to indicate a lesser degree of grammaticalization compared to auxiliaries like *avere* and *essere*, fully integrated in verbal inflection. Yet, they show certain properties of auxiliaries, insofar as they can convey certain features of the infinitive verb, among which auxiliary selection.

This term has been chosen instead of ‘restructuring’ to address the specific phenomenon of auxiliary selection in infinitive constructions. From a morphosyntactic perspective, certain markers related to the infinitive verb may attach to the auxiliary. This is the case not only with

negation or clitics but also with the choice between ‘be’ and ‘have’ as markers of the perfect tense, depending on the agentivity or internal cause traits of the subject of the infinitive verb.

In contrast, no auxiliary variation is possible with infinitive constructions such as *Piero è andato a prendere il latte* ‘Piero went to get the milk’ (see Rizzi 1976a) since *andare* ‘go’ is not a semi-auxiliary, but rather a main verb selecting invariable ‘be’, while the infinitive is a subordinate verb.

Put differently, semi-auxiliary verbs such as *iniziare*, *cominciare*, *continuare* and *finire* allow the infinitive to have a significant role in selecting the auxiliary, while main verbs such as *andare* don’t.

Some of the verbs in the corpus of this study have been retrieved among the results of the following CQL:

(139) [tag= “NOUN|PRO: pers”] [lemma= “avere”] [tag= “VER: ppast”].

Other occurrences were added with a specific CQL:

(140) [tag= “NOUN|PRO: pers”] [lemma= “avere”][word= “finito|finita|finiti|finite”][tag= “PRE”][tag= “VER:infi”]

The semantic factors analyzed in simple verbs were considered here, too: human trait, animacy, agentivity, and internal cause. In terms of syntactic parameters, however, four types of infinitives were distinguished: transitive verbs, intransitive verbs selecting ‘have’, intransitive verbs selecting ‘be’, and finally, intransitive verbs allowing auxiliary alternation (Table 3.49).

Semantic factors	Syntactic factors
Subject [+Human]	V AUX + V INF Transitive (auxiliary ‘have’)
Subject [+Animate]	V AUX + V INF Intransitive (auxiliary ‘have’)
Subject [+Agentive]	V AUX + V INF Intransitive (auxiliary ‘be’)
Subject [+Internal Cause]	V AUX + V INF Intransitive (auxiliary ‘have’   ‘be’)

TABLE 3.49 - Parameters analyzed for semi-auxiliaries.

In the following pages, I will present the output of the corpus analysis of the four semi-auxiliaries verbs found in the corpus by underlining the significant role played by the infinitive.

### 3.9.1 Iniziare

The verb *iniziare* ‘begin’, ‘start’ in its semi-auxiliary function exhibits 97% of occurrences with ‘have’ and 3% with ‘be’ in the ItTenTen (16) corpus, with 18,911 occurrences of ‘have’ and 680 of ‘be’. Thus, ‘have’ prevails.

**Auxiliary selection properties of the infinitive verb** The type of infinitive verb appears to influence whether *iniziare* has either ‘have’ or ‘be’ as an auxiliary: when the infinitive is a transitive verb, the auxiliary is ‘have’.

Otherwise, the preferred auxiliary is ‘be’, as depicted in Figure 3.34 and Table 3.50.

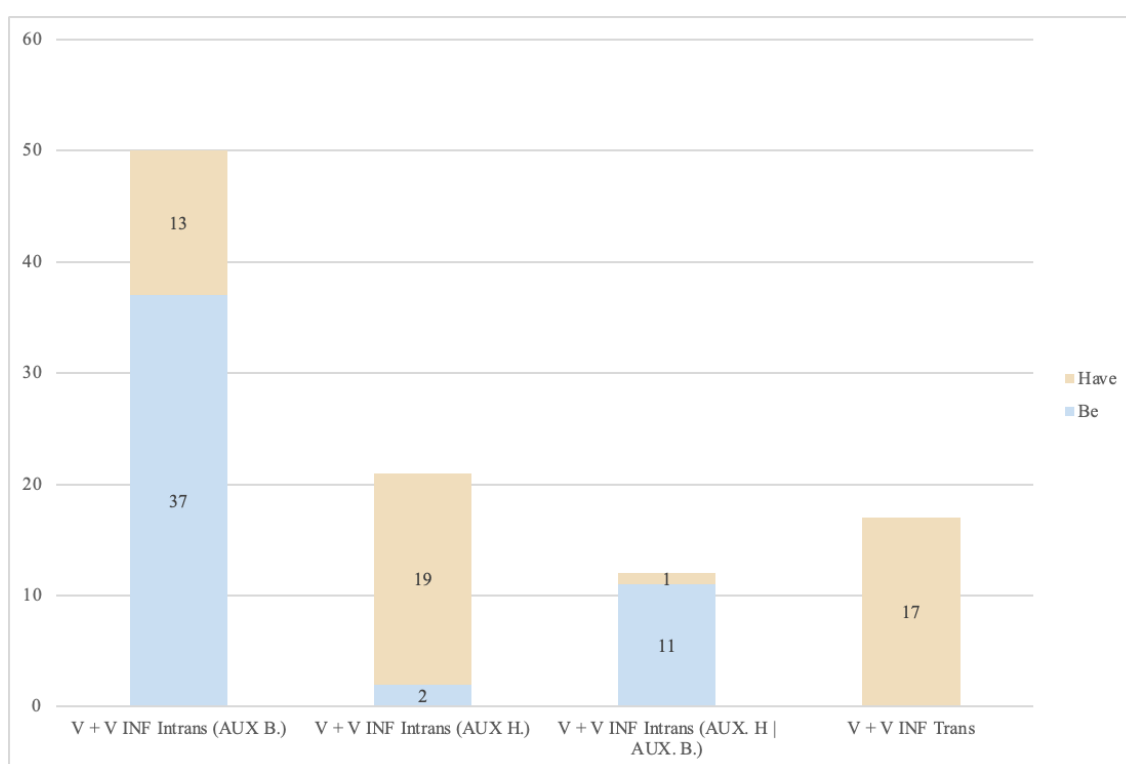


FIGURE 3.34 – Correlation of ‘have’ and ‘be’ with the auxiliary selection of the infinitive verb

Factors	Data ‘Be’	Data ‘Have’	Total
<b>V + V INF Intrans (AUX be)</b>	37	13	50
<b>V + V INF Intrans (AUX have)</b>	2	19	21
<b>V + V INF Intrans (AUX have AUX be)</b>	11	1	12
<b>V + V INF Trans</b>	/	17	50
<b>Total</b>	50	50	100

TABLE 3.50 – Frequency of ‘have’ and ‘be’ based on the auxiliary selection of the infinitive verb.

When the infinitive is transitive (140) or a ‘have’ selecting intransitive verb (141), the semi-auxiliary verb *iniziare* mostly occurs with ‘have’:

(140) [...]io qualche sogno ho iniziato a realizzarlo [...].  
 I some dream have.PRS.1SG start.PTCP to realize.INF.it

‘[...] I have started to realize some dreams [...].’

(141) [...] l'altra sera il mio cane ha iniziato a starnutire.  
 DEF.ART.last night DEF.ART my dog have.PRS.3SG start.PTCP to sneeze.INF

‘[...] last night my dog started sneezing.’

When the infinitive verb allows both ‘have’ and ‘be’ as an auxiliary, as is the case for *circulare*, *piovere*, the perfect tense of *iniziare* contains ‘be’ as an auxiliary (142; 143), with the exception of a single occurrence with ‘have’ (144):

(142) La sera è iniziato a piovere [...].  
 DEF.ART evening be.PRS.3SG start.PTCP to rain.INF

‘In the evening it started to rain [...].’

(143) Le dichiarazioni del primo cittadino sono iniziate a  
 DEF.ART statements of.DEF.ART first citizen be.PRS.3PL start.PTCP.F.PL to

circolare soltanto ieri [...].  
 circulate.INF only yesterday

‘The first citizen’s statements began to circulate only yesterday [...].’

(144) Il video ha iniziato a circolare parecchio sui  
 DEF.ART video hav.-PRS.3SG start.PTCP to circulate.INF a lot on.DEF.ART

social media cinesi [...].  
 social media chinese

‘The video started circulating a lot on Chinese social media [...].’

The sentences in (143) and (144) feature the same verb (*circolare* ‘circulate’) and similar subjects, thus confirming that both auxiliaries can be used when the infinitive verb allows for both. However, ‘be’ is more common.

When the infinitive is a ‘be’ selecting verb, variation occurs, but again ‘be’ is more frequent: in 37 data occurrences ‘be’ is selected like in (145), in 13 occurrences ‘have’ is selected like in (146):

(145) [...] l'effetto degli antidolorifici è iniziato a diminuire.  
 DEF.ART.effect of DEF.ART painkillers be.PRS.3SG begin.PTCP to diminish.INF  
 '[...] the effect of painkillers began to diminish.'

(146) La frequenza della malattia ha iniziato a decrescere  
 DEF.ART frequency of DEF.ART disease have.PRS.3SG start.PTCP to decrease.INF  
 dal 1995 [...].  
 since DEF.ART 1995  
 'The frequency of the disease has started to decrease since 1995 [...].'

As demonstrated in (146), while 'have' is possible with 'be' selection infinitive verbs with a non-human subject, 'have' is predominant with human subjects, often characterized by agentivity, or with internal cause subjects.

**Human and non-human subjects** The feature 'human' remains crucial even with semi-auxiliary verbs, exhibiting a correlation with the auxiliary of 'have' for infinitive verbs allowing both the auxiliaries 'have' and 'be'.

Conversely, non-human subjects typically favor 'be' only when the infinitive is a 'be' selection verb as well as when the infinitive verb allows both auxiliaries, as illustrated in Figure 3.35 and Table 3.51.

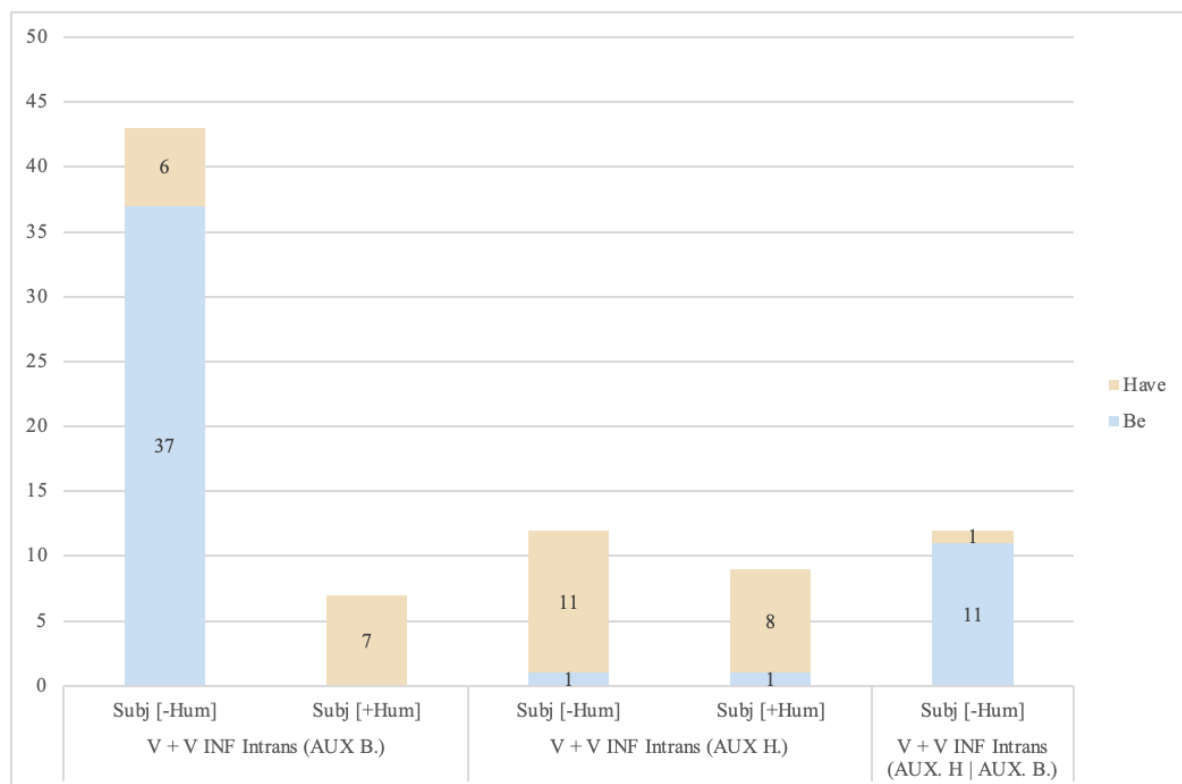


FIGURE 3.35 – Correlation of ‘have’ and ‘be’ with the different types of intransitives and the human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>V + V INF Intrans (AUX B.) Subj [-Hum]</b>	37	6	43
<b>V + V INF Intrans (AUX B.) Subj [+Hum]</b>	/	7	7
<b>V + V INF Intrans (AUX H.) Subj [-Hum]</b>	1	11	12
<b>V + V INF Intrans (AUX H.) Subj [+Hum]</b>	1	8	9
<b>V + V INF Intrans (AUX H  AUX B) Subj [+Hum]</b>	11	1	12
<b>Total</b>	50	33	83

TABLE 3.51 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives and the human/non-human subjects.

The auxiliary selection properties of the infinitive verb are the first relevant factor: when the infinitive is a ‘have’ selecting verb, *iniziare* occurs mostly with ‘have’, independently of the type of subject.

However, human trait is also very relevant: when the subject is human, the auxiliary is ‘have’ with both ‘have’ selecting infinitive verbs like in (147) and ‘be’ selecting infinitive verbs like in (148):

(147) Ed io ho iniziato a bere e fumare fino a  
 and I have.PRS.1SG start.PTCP to drink.INF and smoke.INF until

ridurmi ad uno straccio.  
 reduce.REFL-INF to INDF.ART rag

‘And I started drinking and smoking until I was reduced to a rag.’

(148) [...] da dopo un anno dalla scoperta della sua malattia  
 from after INDF.ART year from.DEF.ART discovery of.DEF.ART her illness

abbiamo iniziato ad andare da lei.  
 have.PRS.1PL start.PTCP to go.INF to her

‘[...] from a year after the discovery of her illness we started to go to her.’

Non-human subjects are associated with the auxiliary ‘have’ when the infinitive verb selects ‘have’ like in (149), and with ‘be’ when the infinitive verb selects ‘be’ like in (150). In the case of the ‘be’ selecting infinitive verb, ‘have’ occasionally occurs as well, see (151).

(149) [...] i suoi occhi e le orecchie hanno iniziato a sanguinare.  
 DEF.ART his eyes and DEF.ART ears have.PRS.3PL start.PTCP to bleed.INF

‘[...] his eyes and ears began to bleed.’

(150) [...] le cose sono iniziate ad andare male.  
 DEF.ART things be.PRS.3PL start.PTCP.3PL to go.INF wrong

‘[...] things started to go wrong.’

(151) [...] la curva della crescita ha iniziato a calare verso  
 DEF.ART curve of.DEF.ART growth have.PRS.3SG start.PTCP to drop.INF towards

il basso da febbraio scorso.  
 DEF.ART down since February last

‘[...] the growth curve has been dropping downward since last February.’

Therefore, two impactful parameters have been observed thus far: the auxiliary selection properties of the infinitive and the human subject. A ‘have’ selecting infinitive verb exhibits a strong influence, leading to the selection of ‘have’ with non-human subjects. Conversely, when the subject is human, this factor is dominant enough to co-occur with ‘have’ even with ‘be’ selecting verbs.

**Agentivity, lack of agentivity and internal cause** When the infinitive is a ‘have’ selecting verb, the preferred auxiliary is ‘have’, both with internal cause subjects and non-agentive subjects. Conversely, when the infinitive is a ‘be’ selecting verb, both auxiliaries occur when the subject is internal cause; when the subject is non-agentive, the preferred auxiliary is ‘be’, as shown in Figure 3.36 and Table 3.52.

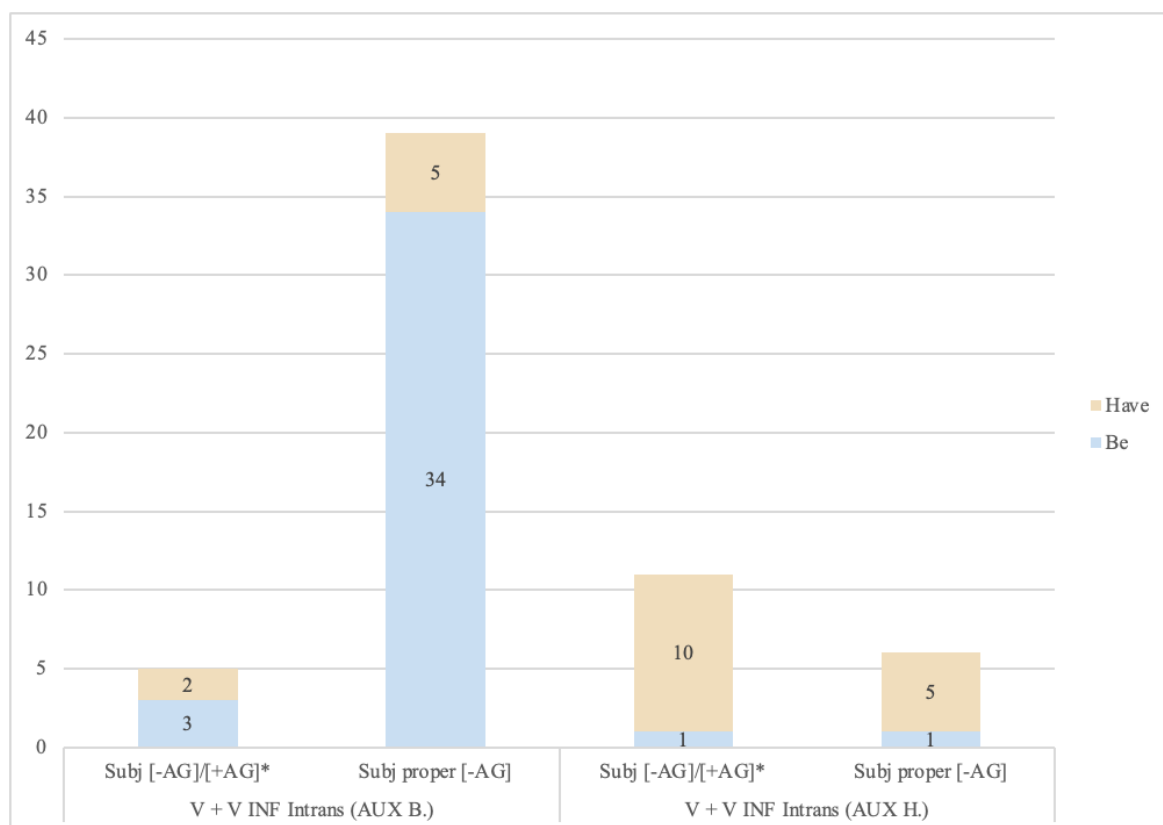


FIGURE 3.36 – Correlation of ‘have’ and ‘be’ with the different types of intransitives and the agentive/non-agentive/internal cause subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
V + V INF Intrans (AUX B.) Subj [-AG]/[+AG] *	3	2	5
V + V INF Intrans (AUX B.) Subj proper [-AG]	34	5	39
V + V INF Intrans (AUX H.) Subj [-AG]/[+AG] *	1	10	11
V + V INF Intrans (AUX H.) Subj proper [-AG]	1	5	6
<b>Total</b>	<b>39</b>	<b>22</b>	<b>61</b>

TABLE 3.52 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives and agentive/non-agentive/internal cause subjects.

In Figure 3.36 and Table 3.52, the focus is on the ‘have’ and ‘be’ selecting intransitive verbs since the transitive verbs only present ‘have’. Noteworthy are the ‘be’ selecting verbs when combined with an internal cause subject. In such cases, there is significant variation, with both ‘have’ (152) and ‘be’ (153).

(152) La nebbia ha iniziato ad alzarsi già da un po'.  
 DEF.ART fog have.PRS.3SG start.PTCP to lift.REFL.INF already from a.bit

‘The fog has been lifting for some time now [...].’

(153) Finalmente alle 13:30 il vento è iniziato ad  
 Finally at.DEF.ART 1:30 p.m. DEF.ART wind be.PRS.3SG start.PTCP to

entrare[...].  
 come.in.INF

‘Finally, at 1:30 p.m. the wind started to come in [...].’

The non-agentive subjects are all non-human. The preferred auxiliary in these cases is ‘be’ (154), but the use of ‘have’ (155) is interesting because it could suggest that ‘have’ is emerging as the default auxiliary in infinitive constructions with semi-auxiliaries of aspect.

(154) [...] la mortalità è iniziata a scendere sensibilmente.  
 DEF.ART mortality be.PRS.3SG begin.PTCP.F.SG to fall.INF significantly

‘[...] mortality began to fall significantly.’

(155) Dopo alcuni giri, [...] le gomme hanno iniziato a  
 after some laps DEF.ART wheels have.PRS.3PL begin.PTCP to

calare [...]  
 drop.INF

‘After a few laps [...], the wheels began to drop [...].’

Besides (154), in most occurrences with non-human non-agentive subjects, the auxiliary is ‘be,’ but 5 occurrences in the corpus out of 39 demonstrate that ‘have’ is also possible.

In conclusion, the infinitive verb plays a significant role in auxiliary selection. When the infinitive is a ‘have’ selecting verb, both transitive and intransitive, ‘have’ is strongly favored.

Conversely, when the infinitive is a ‘be’ selecting verb, ‘be’ is preferred. With ‘be’ selecting infinitive verbs, there is more variation. Human trait and agentivity influence the selection: even with ‘be’ selecting infinitive verbs, human and agentive subjects only occur with ‘have’. The internal cause subject allows for a lot of variation, while non-human (and non-agentive) subjects strongly correlate with ‘be’.

This choice may be due to the progressive grammaticalization of ‘have’ as the default auxiliary, a trend observed in French (Rea 2018), and which has eventually become dominant in other Romance languages such as Spanish (Aranovich 2003). French shows signs of the expansion of ‘have’ even for verbs that traditionally would have selected ‘be’ (Rea 2018: 2).<sup>18</sup>

In Spanish, ‘have’ has throughout its history completely replaced ‘be’, as demonstrated by Aranovich, who states that “predicates that have a more patient-like subject are the last ones to lose their ability to select *ser*” (Aranovich 2003: 1). Italian semi-auxiliaries tend also to preserve ‘be’ selection when the subject is non-agentive and non-human.

Table 3.53 illustrates the most predominant co-occurring factors for the choice of auxiliaries with *iniziare*.

Infinitive	Agentivity	Human Trait	Auxiliary
‘Have’ selecting infinitive verb	Agentive, Internal	Human, Non-human →	Have
	Cause, Non-agentive		
‘Be’ selecting infinitive verb	Non-agentive	Non-human →	Be
	Internal cause	Human/Non-human →	Have/Be

TABLE 3.53 – Primary factors affecting the selection of the auxiliary in *iniziare*.

The most important factors that emerge from the analysis are, therefore, the auxiliary selection properties of the infinitive verb (‘have’ selecting or ‘be’ selecting), followed by the human trait of the subjects, which often co-occurs with agentivity.

### 3.9.2 *Cominciare*

*Cominciare* used as a semi-auxiliary shows 97% occurrences with ‘have’ and 3% with ‘be’ in the ItTenTen (16) corpus, with 15,035 occurrences of ‘have’ and 530 of ‘be’. Like *iniziare*, the CQL search found more occurrences of *cominciare* with ‘have’ than with ‘be’.

***Auxiliary selection properties of the infinitive verb*** When the infinitive is a ‘have’ selecting verb, either transitive or intransitive, only ‘have’ is selected. When the infinitive is a verb that allow both auxiliaries, the preferred one is ‘be’ (although there is only a small amount of data).

<sup>18</sup> “Plusieurs études suggèrent qu’un peu partout dans la francophonie des locuteurs de langue maternelle emploient à l’oral les deux auxiliaires, *avoir* et *être*, avec des verbes qui, dans la langue normative, ne peuvent être utilisés qu’avec être (ex. *J’ai tombé*). Ce phénomène serait signe qu’une généralisation de l’auxiliaire *avoir* a cours dans ces régions francophones” (Rea 2018 :2).

Finally, when the infinitive is a ‘be’ selecting verb, both auxiliaries occur, with ‘be’ being more frequently used. This picture is illustrated in Figure 3.37 and Table 3.54.

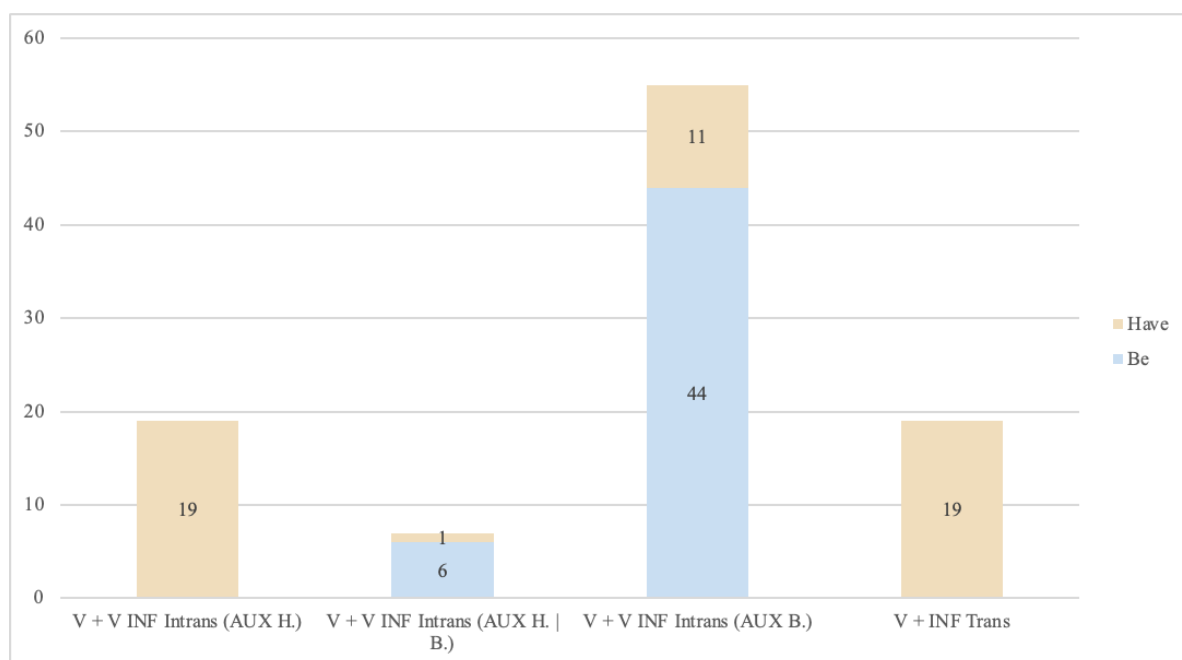


FIGURE 3.37 – Correlation of ‘have’ and ‘be’ with the different types of intransitives.

Factors	Data ‘Be’	Data ‘Have’	Total
V + V INF Intrans (AUX H.)	/	19	19
V + V INF Intrans (AUX H.   B.)	6	1	7
V + V INF Intrans (AUX B.)	44	11	55
V + V INF Trans	/	19	19
<b>Total</b>	50	50	100

TABLE 3.54 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives.

As depicted in Figure 3.37 and Table 3.53, the auxiliary ‘have’ is possible in all contexts. There are 19 occurrences of ‘have’ selecting intransitive verbs like in (156) and 19 occurrences with transitives like in (157).

(156) La mia mente ha cominciato a viaggiare.  
DEF.ART my mind have.PRS.3SG begin.PTCP to travel.INF

‘My mind began to travel.’

(157) Una volta a casa ho cominciato a leggerlo [...].  
INDF.ART time at home have.PRS.1SG begin.PTCP to read.INF.it

‘Once home, I started reading it [...].’

When infinitive verbs permit both auxiliaries, the preferred auxiliary is ‘have’ like in (158), although ‘be’ (159) remains acceptable in the same context.

- (158) Verso mezzogiorno ha cominciato a nevicare [...].  
 Around noon have.PRS.3SG begin.PTCP to snow.INF  
 ‘Around noon it began to snow [...].’

- (159) Nel frattempo, è cominciato a piovere [...].  
 in.DEF.ART meantime be.PRS.3SG begin.PTCP to rain.INF  
 ‘In the meantime, it began to rain [...].’

The sentences in (158) and (159) depict atmospheric events, i.e. verbs known for exhibiting auxiliary variation (Sorace 2000: 878): “weather verbs display apparent free variation in auxiliary selection in Italian”. With the exception of two occurrences (involving the verbs *sbiadire* ‘fade’ and *lampeggiare* ‘flash’), all infinitive verbs with ‘have-be’ auxiliary alternation denote atmospheric events.

‘Be’ selecting infinitive verbs prefer ‘be’ as an auxiliary for *cominciare* (160), but admit also ‘have’, with only 11 instances presenting ‘have’ like in (161) compared to 44 instances with ‘be’.

As already noticed with *iniziare*, ‘be’ selecting infinitive verbs allow auxiliary variation for the semi-auxiliary.

- (160) [...] la situazione è cominciata a precipitare [...].  
 DEF.ART situation be.PRS.3SG begin.PTCP.F.SG to precipitate.INF  
 ‘[...] the situation began to precipitate [...].’

- (161) [...] precariato e disoccupazione hanno cominciato a dilagare [...].  
 precariousness and unemployment have.PRS.3PL begin.PTCP to overflow.INF  
 ‘[...] precariousness and unemployment began to overflow [...].’

The examples in (160) and (161) represent two similar configurations in which both auxiliaries seem acceptable. Both involve non-human subjects, once again highlighting the influence of the subject type on auxiliary selection. The first factor of influence is the infinitive verb and its auxiliary selection properties, the second factor is the human trait, with human subjects favoring ‘have’ when the auxiliary’s variation is possible, particularly with ‘be’ selecting verbs.

**Human and non-human subjects** The role of the infinitive is clear: when it is a ‘have’ selecting verb, it invariably co-occurs with ‘have’, regardless of whether the subject is human or non-human. Conversely, when the infinitive pairs with a verb that tends to select ‘be’ as its auxiliary, the distribution of the two auxiliaries becomes variable.

This variability is influenced by the human trait: with human subjects, the auxiliary ‘have’ predominates, while with non-human subjects, ‘be’ emerges as the preferred auxiliary. This phenomenon is shown in Figure 3.38 and in Table 3.55.

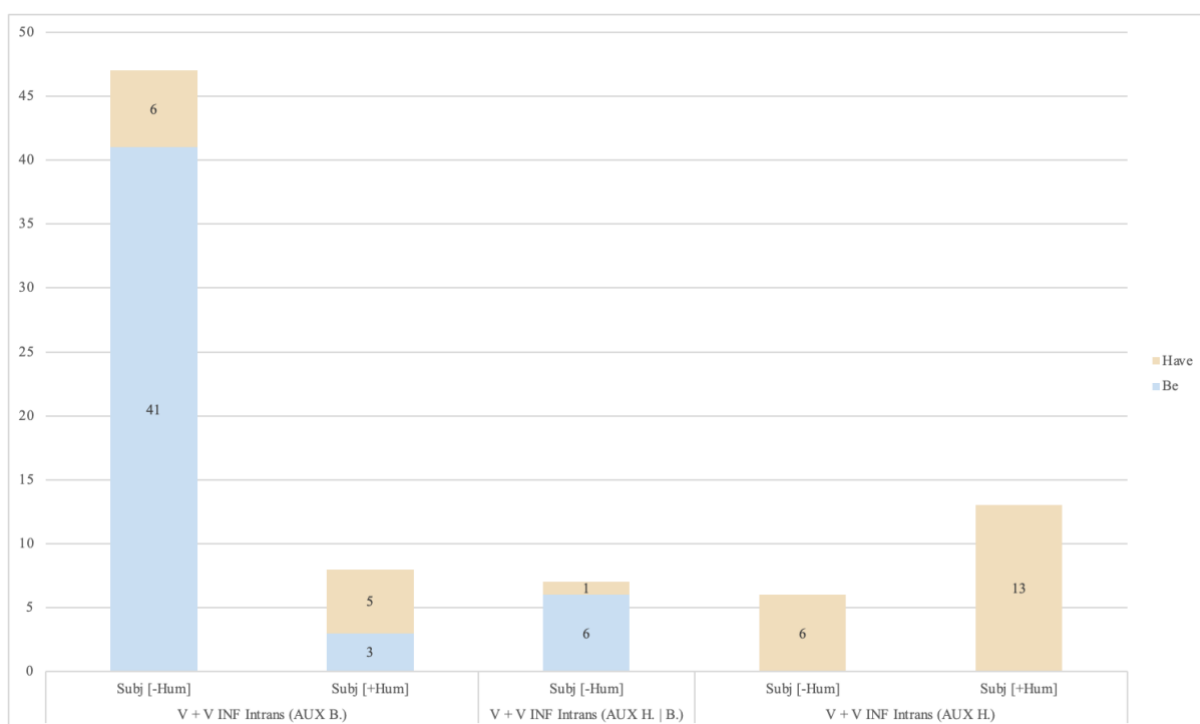


FIGURE 3.38 – Correlation of ‘have’ and ‘be’ with the different types of intransitives and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
V + V INF Intrans (AUX B.) Subj [-Hum]	41	6	47
V + V INF Intrans (AUX B.) Subj [+Hum]	3	5	8
V + V INF Intrans (AUX H.   B.) Subj [-Hum]	6	1	7
V + V INF Intrans (AUX H.) Subj [-Hum]	/	6	6
V + V INF Intrans (AUX H.) Subj [+Hum]	/	13	13
<b>Total</b>	<b>50</b>	<b>49</b>	<b>93</b>

TABLE 3.55 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives and human/non-human subjects.

As depicted in Figure 3.38 and Table 3.55, instances involving a transitive verb as the infinitive are not presented, since they select invariably the auxiliary ‘have’ and combine systematically with a human subject.

It is crucial to note that the feature ‘human’ does not directly influence the selection of the auxiliary when the infinitive aligns with a verb favoring ‘have’. In such cases, the auxiliary ‘have’ is required by the nature of the infinitive itself.

However, in contexts where variation in auxiliary selection is possible, the feature ‘human’ becomes relevant. Specifically, when the subject is non-human, ‘be’ emerges as the preferred auxiliary (162), while ‘have’ may also be possible (163).

- (162) [...] anche i risultati sono cominciati ad arrivare.  
 even DEF.ART results be.PRS.3PL begin.PTCP.M.PL to come.INF  
 ‘[...] results also become to come.’

- (163) [...] anche le macchie sulla pelle hanno cominciato a  
 even DEF.ART spots on.DEF.ART skin have.PRS.3PL begin.PTCP to  
 essere meno evidenti.  
 be.INF less noticeable.  
 ‘[...] even the spots on the skin began to be less noticeable.’

Conversely, when the subject is human and the infinitive is associated with a ‘be’ selecting verb, it is possible for the auxiliary ‘be’ to be selected (165), but the preferred auxiliary in such cases is typically ‘have’ (164).

- (164) [...] da quest’anno ho cominciato ad allenarmi [...].  
 since this.year have.PRS.1SG start.PTCP to train.REFL.INF  
 ‘[...] since this year, I started training [...].’

- (165) La gente è cominciata ad arrivare verso le  
 DEF.ART people be.PRS.3SG start.PTCP.F.SG to come around DEF.ART  
 22:10/15.  
 10:10/15 p.m.  
 ‘People started arriving at around 10:10/15 p.m.’

While not all instances of ‘have’ with ‘be’ selecting verbs necessarily feature human and agentive subjects, the presence of agentivity and human traits tends to increase the prevalence of ‘have’.

*Agentivity, lack of agentivity and internal cause* ‘Have’ selecting infinitive verbs predominantly involve internal cause and agentive subjects and *cominciare* invariably occurs

with ‘have’. Conversely, with ‘be’ selecting infinitive verbs, *cominciare* occurs with both ‘have’ and ‘be’. A significant proportion of occurrences contains non-agentive subjects, which exhibit a preference for ‘be’ as the auxiliary.

Although infrequent, instances with agentive subjects (which also are human) demonstrate a preference for ‘have’. Internal cause subjects exhibit a mixed pattern of both ‘have’ and ‘be’ auxiliaries. These observations are represented in detail in Figure 3.39 and Table 3.56.

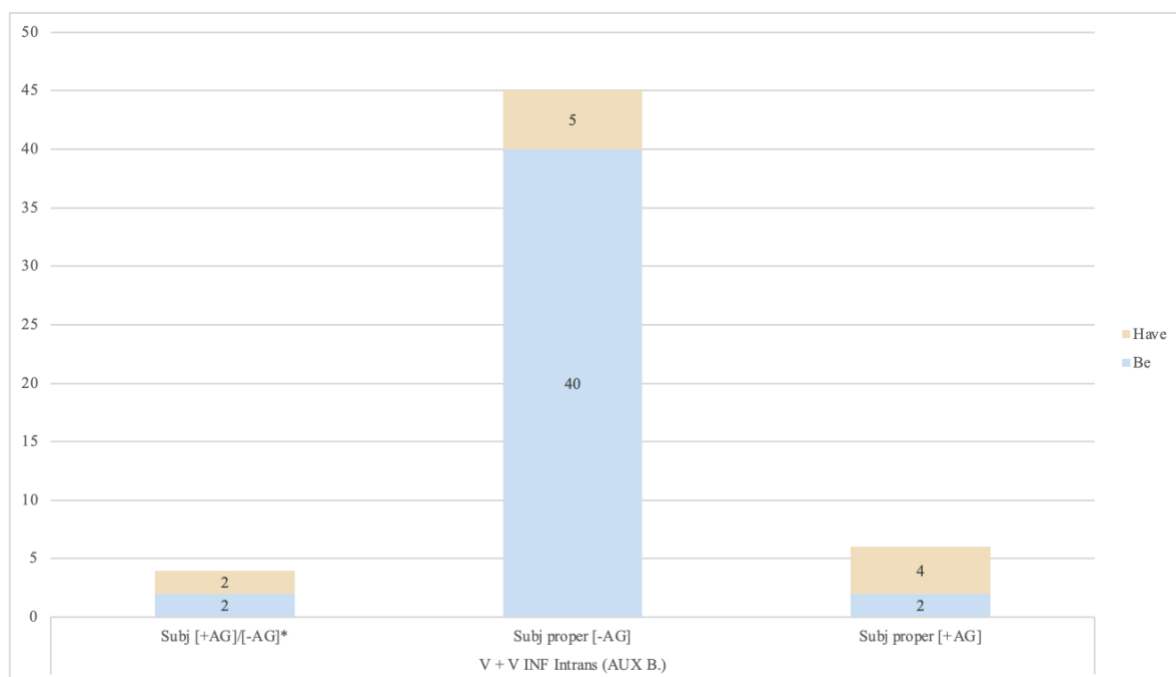


FIGURE 3.39 – Correlation of ‘have’ and ‘be’ with the different types of intransitives and agentive/internal cause/non-agentive subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>V + V INF Intrans (AUX B.) Subj [-AG/ [+AG] *</b>	2	2	4
<b>V + V INF Intrans (AUX B.) Subj proper [-AG]</b>	40	5	45
<b>V + V INF Intrans (AUX B.) Subj proper [+AG]</b>	2	4	6
<b>Total</b>	44	11	55

TABLE 3.56 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives and agentive/internal cause/non-agentive subjects.

As indicated by both the figure and the table, the predominant occurrence of ‘be’ with ‘be’ selecting infinitives is observed in contexts involving non-agentive subjects that are typically also non-human. There is a clear preference for ‘be’ like in (166), although ‘have’ is not uncommon, as illustrated in (167):

(166) [...] da lì la situazione è cominciata a  
 from there DEF.ART situation be.PRS.3SG begin.PTCP.F.SG to

migliorare [...].  
 improve. INF

‘[...] from there the situation began to improve [...].’

(167) [...] ogni cosa ha cominciato poi ad essere più personale  
 every thing have.PRS.3SG begin.PTCP then to be.INF more personal

e distorta [...].  
 and distorted

‘[...] ogni cosa ha cominciato poi ad essere più personale e distorta [...].’

Even if the non-agentive subjects are mostly non-human, human subjects do occur (168):

(168) [...] dato che il suo pubblico principale fatto di  
 since that DEF.ART its audience main make.PTCP of

pensionati ha cominciato [...] a morire.  
 retirees have.PRS.3SG begin.PTCP to die.INF

‘[...] since its main audience made up of retirees has begun [...] to die.’

(168) is the only instance offering a non-agentive subject that is human. All other occurrences of non-agentive subjects are non-human, whether ‘have’ or ‘be’ is used. The presence of ‘have’ with this type of subject, despite limited evidence, suggests again that ‘have’ is the default auxiliary in infinitive constructions with aspectual semi-auxiliaries.

The proper agentive subjects, on the other hand, are exclusively human and predominantly favor ‘have’ as the auxiliary. The preference for ‘have’ is exemplified in (169), while (170) shows that ‘be’ with the same agentive subject is nevertheless possible:

(169) Negli ultimi anni molti artisti hanno cominciato ad  
 in.DEF.ART recent years many artists have.PRS.3PL begin.PTCP to

appropriarsi di questi materiali [...].  
 appropriate.REFL. INF of these materials

‘In recent years many artists have begun to appropriate these materials [...].’

(170) [...] le persone sono cominciate ad arrivare intorno alle  
DEF.ART people be.PRS.3PL start.PTCP to come.PTCP around at.DEF.ART

15:30.  
 3:30p.m.

‘[...] people started arriving around 3:30 p.m.’

(169) presents a reflexive verb that always selects ‘be’ in Italian. As Sorace (2000: 861) states, “all verbs marked with the clitic *si/se* select auxiliary ‘be’ in Italian.” In these examples, the ‘be’ selecting verb, the human trait reflecting the inherent qualities of the subject, and the agentive trait indicating the subject’s ability to intentionally perform actions, all play pivotal roles.

While the ‘be’ selecting infinitive verb may tend to favor the auxiliary ‘be’ for *cominciare*, the human and agentive traits of the subjects can counterbalance this tendency, leading to the selection of ‘have’.

When the subject is an internal cause, both auxiliaries occur in the sample, with two occurrences of ‘have’, as in (171), and two of ‘be’, as in (172):

(171) [...] un'altra idea ha cominciato a formarsi.  
INDEF.ART.other idea have.PRS.3SG start.PTCP to form.INF

‘[...] un'altra idea ha cominciato a formarsi.’

(172) La produzione è cominciata a calare [...] l'attività  
DEF.ART production be.PRS.3SG begin.PTCP.F.SG to decline.INF DEF.ART.activity

si è andata del tutto estinguendosi.  
REFL be.PRS.3SG go.PTCP.F.SG completely becoming.extinct.REFL

‘Production began to decline [...] the business became completely extinct.’

In sentences (171) and (172), the subjects are internal causes with a link to a human referent. For instance, in (171), the idea ‘taking shape in a person’s mind’ implies a cognitive process driven by human agency. Similarly, in (172), the decrease in production and the consequent business’ extinction are attributed to the actions of people.

In such scenarios, the causative nature of the subject linked to a human referent can indeed lean towards the usage of ‘have’ as the auxiliary. This is because ‘have’ often aligns with causative constructions, indicating agency or control over process. However, because of the presence of a ‘be’ selecting infinitive verb and the intermediate reading of internal cause, which implies a lack of proper agentivity, ‘be’ is also possible.

In sum, the dynamic interplay between causativity linked to a human referent and the presence of a ‘be’ selecting infinitive verb in such constructions can result in a nuanced choice between ‘have’ and ‘be’ as the auxiliary.

Therefore, the analysis of *cominciare* ‘begin’ ‘start’ confirms the hypotheses already made for the *iniziare*: the infinitive plays a major role, especially when it is a ‘have’ selecting verb, the auxiliary ‘have’ is used for the perfect tense of *cominciare*.

However, additional factors such as the human trait and agentivity also influence auxiliary selection, as they favor the auxiliary ‘have’ even with ‘be’ selecting verbs. The most significant auxiliary alternation occurs with ‘be’ selecting infinitive verbs in the presence of internal cause subjects.

Noteworthy, there is a residue of occurrences of infinitive constructions with a ‘be’ selecting infinitive verb and with non-human and non-agentive subjects and where the auxiliary ‘have’ nonetheless appears. As we will see in the following chapter, the choice of the auxiliary ‘have’ in this syntactic configuration is normal in French. These rare cases in Italian could be an indication that grammaticalization of ‘have’ as the default auxiliary has also started in Italian.

Table 3.57 summarizes the key factors for the selection of ‘have’ and ‘be’ with *cominciare*, providing a comprehensive overview of the nuanced dynamics at play in auxiliary usage.

Infinitive	Agentivity	Human Trait	Auxiliary
‘Have’ selecting	Agentive, Internal Cause, Non-agentive	Human, Non-human →	Have
‘Be’ selecting verb	Non-agentive	Non-human →	Be
	Internal cause	Human/Non-human →	Have/Be

TABLE 3.57 – Primary factors affecting auxiliary selection of *cominciare* as a semi-auxiliary in infinitive constructions.

### 3.9.3 *Continuare*

*Continuare* ‘continue’ exhibits a substantial prevalence of occurrences with the auxiliary ‘have’ in the ItTenTen (16) corpus, amounting to 97% with ‘have’ (7,625 occurrences) and 3% (267 occurrences) with ‘be’, identified using a specific Corpus Query Language (CQL) designed for capturing semi-auxiliary uses. It is interesting to note that the percentages are exactly the same as for *iniziare* and *cominciare*.

However, it is important to acknowledge that the significant predominance of ‘have’ may also stem from the nature of the infinitive associated with *continuar*. Specifically, the infinitive is often a ‘have’ selecting verb, contributing to the considerable prevalence of ‘have’ as the preferred auxiliary in these constructions.

**Auxiliary selection properties of the infinitive verb** As shown in Figure 3.40 and Table 3.58, the pattern of co-occurrence between the ‘have’ selecting infinitive verb and the auxiliary ‘have’ remains consistent, regardless of whether the verb is intransitive or transitive. However, with ‘be’ selecting infinitive verbs, there is a clear preference for the auxiliary ‘be’,

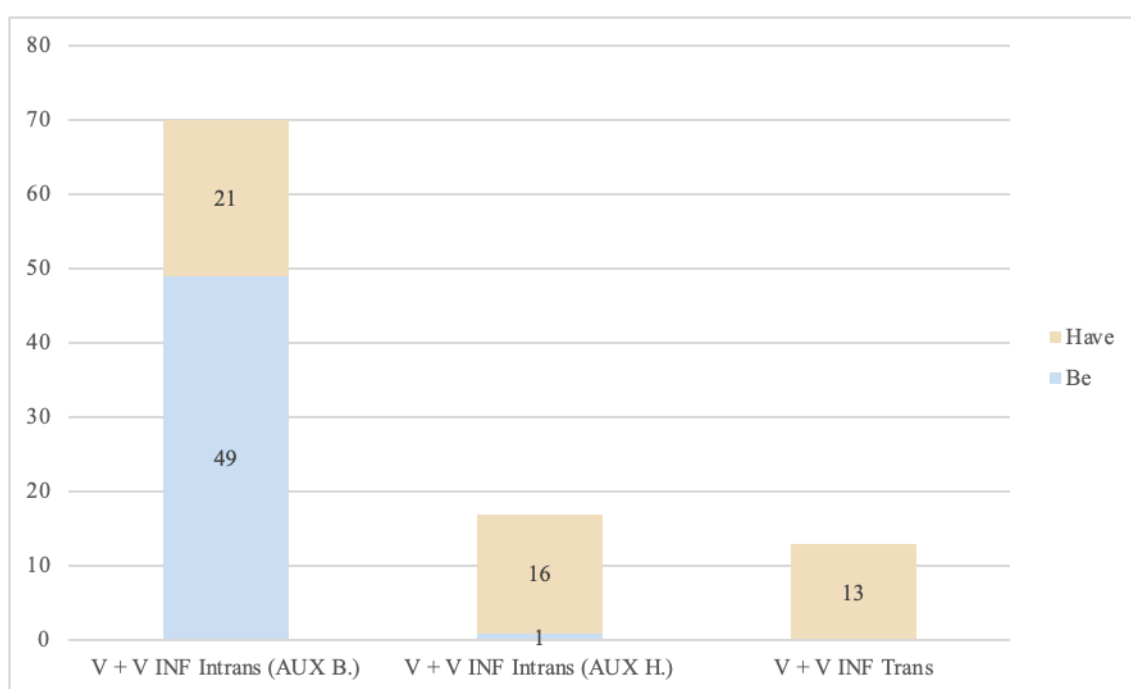


FIGURE 3.40 – Correlation of ‘have’ and ‘be’ with the different types of infinitives.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>V + V INF Intrans (AUX B.)</b>	49	21	70
<b>V + V INF Intrans (AUX H.)</b>	1	16	17
<b>V + V INF Trans</b>	/	13	13
<b>Total</b>	50	50	100

TABLE 3.58 – Frequency of ‘have’ and ‘be’ based on the different types of infinitives.

(173) et (174) provide examples of *continuar* with the auxiliary ‘have’ when the infinitive is a ‘have’ selecting verb, either transitive or intransitive.

(173) [...] il mio datore di lavoro ha continuato a versare  
 DEF.ART my employer of work have.PRS.3SG continue.PTCP to pay.INF

per altri due mesi la rata del prestito.  
 for other two months DEF.ART installment of DEF.ART loan

‘[...] my employer continued to pay the loan installment for another two months.’

(174) In questi vent’anni abbiamo continuato a lavorare [...].  
 In these twenty.years have.PRS.3PL continue.PTCP to work.INF

‘In these twenty years we continued to work [...].’

When the infinitive is a ‘be’ selecting intransitive verb, *continuare* usually occurs with the auxiliary ‘be’ (175):

(175) [...] il numero di omicidi è continuato a salire.  
 DEF.ART number of murders be.PRS.3SG continue.PTCP to rise.INF

‘[...] the number of murders continued to rise.’

(176) [...] questa percentuale ha continuato a crescere.  
 this percentage have.PRS.3SG continue.PTCP to grow.INF

‘[...] this percentage has continued to grow.’

The semi-auxiliary status of *continuare* may lead to the hypothesis that the auxiliary selection properties of the verb in the infinitive are transferred to the semi-auxiliary. This mechanism accounts for the vast majority of cases, encompassing all instances where the verb, transitive or intransitive, is ‘have’ selecting and a majority of cases where the verb, intransitive, is ‘be’ selecting.

Overall, however, the auxiliary ‘have’ predominates in infinitive constructions with an aspectual semi-auxiliary and may appear under the influence of certain semantic features characterizing the subject.

***Human and non-human subjects*** When the infinitive is a verb selecting ‘be’, the human trait of the subject becomes relevant. Human subjects exhibit a balanced distribution of ‘have’ and ‘be’, despite ‘be’ being the expected auxiliary. Conversely, with non-human subjects, there is a preference for ‘be’, as illustrated in Figure 3.41 and Table 3.59.

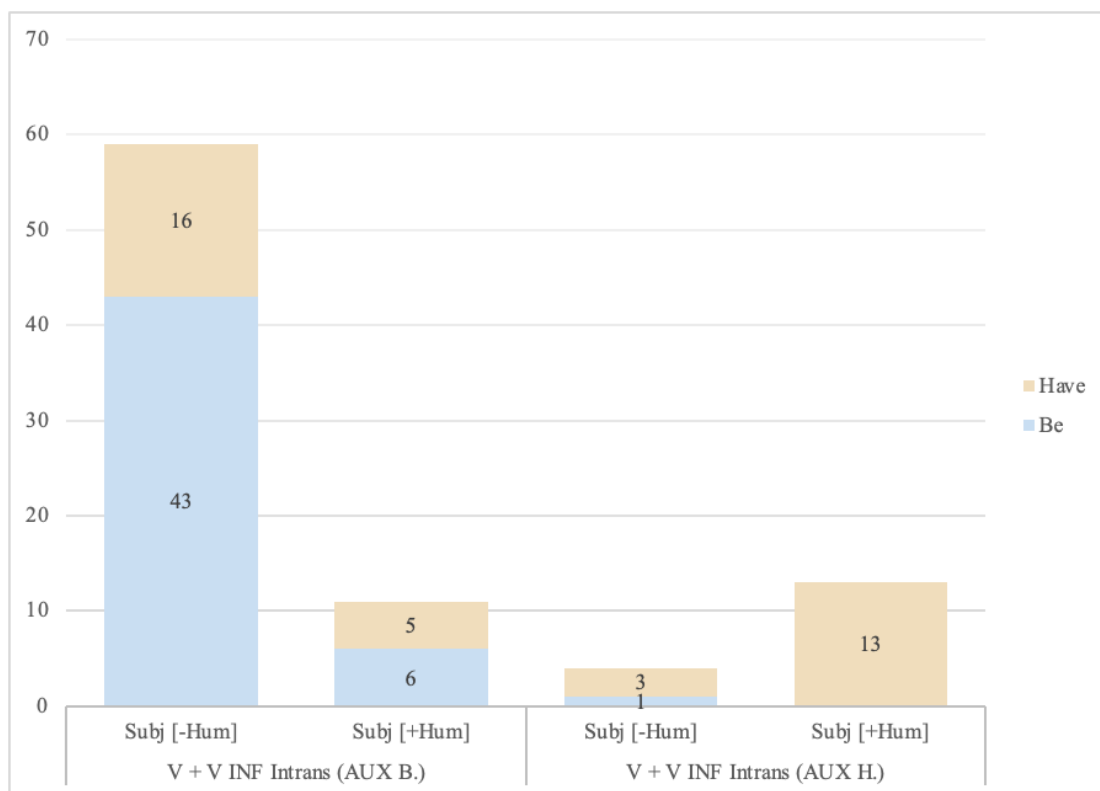


FIGURE 3.41 – Correlation of ‘have’ and ‘be’ with the different types of intransitives and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
V + V INF Intrans (AUX B.) Subj [-Hum]	43	16	59
V + V INF Intrans (AUX B.) Subj [+Hum]	6	5	11
V + V INF Intrans (AUX H.) Subj [-Hum]	1	3	4
V + V INF Intrans (AUX H.) Subj [+Hum]	/	13	13
<b>Total</b>	50	37	87

TABLE 3.59 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives and human/non-human subjects.

Nonetheless, the infinitive has a major impact, given that when it is a verb selecting for ‘have’, the auxiliary remains consistently ‘have’ regardless of whether the subject is human or non-human. But even in this case, the non-human subject can influence auxiliary selection, as the only occurrence of ‘be’ is observed with a non-human subject (e.g., *la terra è continuata a tremare* – ‘the earth has continued to shake’).

Of greater interest are the occurrences with ‘be’. As mentioned earlier, a ‘be’ selecting infinitive verb contributes to the preference for ‘be’. Additionally, the human trait of the subject can also be a relevant factor: when the subject is human, the auxiliary ‘have’ is more frequent, resulting in a 50-50 distribution of each of the two auxiliaries, exemplified in (177) with ‘have’ and (178) with ‘be’:

(177) Ti hanno alzato lo stipendio [...] e tu hai  
 To.you have.PRS.3PL raise.PTCP DEF.ART salary and you have.PRS.2SG

continuato ad essere triste.  
 continue.PTCP to be.INF sad

‘The raised your salary [...] and you continued to be sad.’

(178) Gli abitanti del mio paese sono continuati a  
 DEF.ART inhabitants of.DEF.ART my country be.PRS.3PL continue.PTCP.M.PL to

diminuire.  
 decrease.INF

‘The inhabitants of my country continued to decrease [...].’

Conversely, non-human subjects strongly favor ‘be’ (179) with the ‘be’ selecting intransitive as infinitive.

(179) [...] il prezzo è continuato a salire [...].  
 DEF.ART price be.PRS.3SG continue.PTCP to rise.INF

‘[...] the price continued to rise [...].’

When the infinitive is a ‘be’ selecting verb and the subject is non-human, several factors jointly favor the selection of ‘be’ and account for a strong predominance of this auxiliary. However, there are instances where sentences feature non-human and non-agentive subjects that still select ‘have’, as illustrated in (180) and (181):

(180) L’indice di crescita del mercato ha continuato a  
 DEF.ART.index of growth of.DEF.ART market have.PRS.3SG continue.PTCP to

salire dal ‘98 (+2,8%) [...].  
 rise.INF since.DEF.ART ‘98 (+2,8%)

‘The construction market growth index has continued to rise since ‘98 (+2.8%) [...].’

(181) [...] questa percentuale ha continuato a crescere.  
 this percentage have.PRS.3SG continue.PTCP to grow.INF

‘[...] this percentage has continued to grow.’

In (180) and (181) the selection of ‘have’ is explained in terms of a phenomenon already witnessed: the expansion of ‘have’ at the expense of ‘be’. It could be hypothesized that semi-auxiliary verbs present the process of grammaticalization of ‘have’, shared with other Romance

languages. This is supported by the cooccurrence of many subjects referring to non-human entities with ‘have’.

**Agentivity, internal cause and lack of agentivity** The significant role of the auxiliary selection properties of the infinitive verb, alongside the distinction between human and non-human subjects, is reinforced when examining the correlation with agentivity. Infinitive that are ‘have’ selecting verbs predominantly occur with agentive and internal cause subjects.

Conversely, even when the infinitive is a verb selecting for ‘be’, agentivity and internal cause tend to correlate with the use of ‘have’, while the absence of agentivity aligns with the use of ‘be’. This trend is illustrated in Figure 3.42 and Table 3.60.

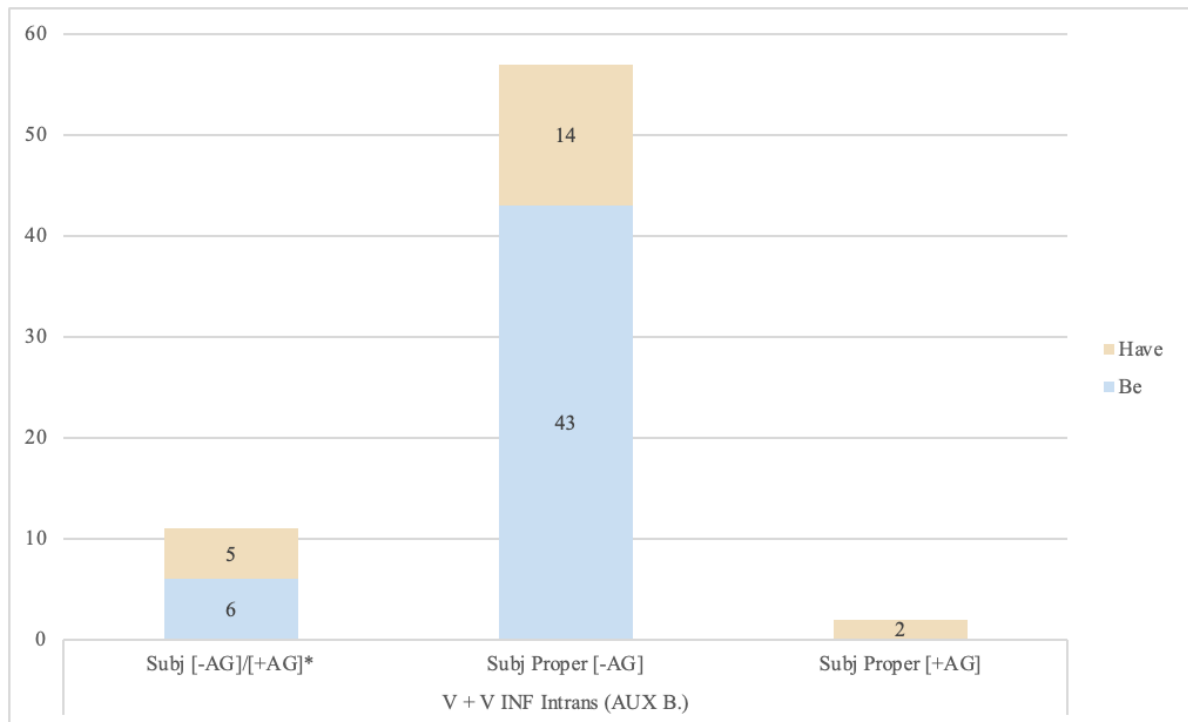


FIGURE 3.42 – Correlation of ‘have’ and ‘be’ with the different types of intransitives and agentive/internal cause/non-agentive subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
V + V INF Intrans (AUX B.) Subj [-AG] / [+AG] *	6	5	11
V + V INF Intrans (AUX B.) Subj proper [-AG]	43	14	57
V + V INF Intrans (AUX B.) Subj proper [+AG]	/	2	2
<b>Total</b>	49	21	70

TABLE 3.60 – Frequency of ‘have’ and ‘be’ based on the different types of intransitives and agentive/internal cause/non-agentive subjects.

Figure 3.42 and Table 3.60 exclusively portray instances involving ‘be’ selecting verbs, as they exhibit the most pronounced differences and variations as to auxiliary selection of *continuare*.

Furthermore, they account for the majority of the dataset, constituting 70 out of 100 occurrences. Within the dataset featuring agentive subjects, instances such as (182) are observed:

- (182) [...] le persone hanno continuato a partire e [...] sono  
 DEF.ART people have.PRS.3PL continue.PTCP to leave.INF and be.PRS.3PL  
 morte in mare.  
 die.PTCP.F.PL in sea  
 ‘[...] people continued to leave [...] and have died at sea.’

In (182), individuals make a conscious decision to leave, demonstrating their control over the action. The presence of a human subject, coupled with this agency leads to the selection of ‘have’. In other words, ‘have’ is favored by the semantic parameters of humanity and the agentivity of the subject, although the infinitive is a ‘be’ selecting verb.

Most variation when the infinitive is a ‘be’ selecting verb occurs with internal cause subjects. Compare (183) with ‘have’ and (184) with ‘be’:

- (183) [...] il fiumiciattolo ha continuato a scorrere.  
 DEF.ART little.river have.PRS.3SG continue.PTCP to flow.INF  
 ‘[...] the little river continued to flow.’

- (184) [...] il petrolio è continuato a sgorgare copioso dopo il  
 DEF.ART oil be.PRS.3SG continue.PTCP to flow.INF copiously after DEF.ART  
 1977.  
 1977  
 ‘[...] oil continued to flow copiously after 1977.’

*Fiumiciattolo* ‘little river’ and *petrolio* ‘oil’ are regarded as internal causes because they are responsible for the process described by the verb, such as the flow, although devoid of volition, and they are affected by this process. In this case, the subject causes the process and undergoes the process, which accounts for the possibility of both ‘have’ and ‘be’.

Moreover, non-agentive subjects with ‘be’ verbs predominantly select ‘be’, even if instances of ‘have’ occur. As mentioned earlier, this extension of ‘have’ may be attributed to the grammaticalization of ‘have’ and its role as a default auxiliary.

In sum: when *continuare* functions as a semi-auxiliary, various factors influence auxiliary selection. Firstly, the auxiliary selection properties of the verb used in the infinitive form have

a major impact: a ‘have’ selecting verb (whether transitive or intransitive) used as the infinitive predominantly correlates with the choice of ‘have’, with little exceptions. In the same vein, when the infinitive is a ‘be’ selecting verb, the auxiliary ‘be’ is predominant, but some other factors may lead to the selection of ‘have’.

In particular, when the subject is human and/or agentive the auxiliary tends to be ‘have’. When the subject is an internal cause, significant variation may occur, because of the intermediate status of internal cause subjects. When the subject is non-human and non-agentive, the auxiliary is typically ‘be’.

Table 3.61 provides a summary of the primary co-occurring factors influencing auxiliary selection within *continuare*.

Infinitive	Agentivity	Human Trait	Auxiliary
‘Have’ selecting	Agentive, Internal Cause, Non-agentive	Human, Non-human →	Have
‘Be’ selecting verb	Non-agentive	Non-human →	Be
	Internal cause	Human/Non-human →	Have/Be

TABLE 3.61 – Primary factors affecting the selection of the auxiliary in *continuare*.

Our observations are globally in line with what has been observed for the semi-auxiliaries *iniziare* and *cominciare*. However, in comparison with the other verbs analyzed thus far, there are more instances with non-agentive subjects selecting ‘have’, even though ‘be’ remains the preferred choice.

As a global tendency, the auxiliary ‘have’ is dominant in infinitive constructions with aspectual semi-auxiliaries in that an infinitive corresponding to a ‘have selecting verb leads to the use of the auxiliary ‘have’ for all aspectual semi-auxiliaries in infinitive constructions, whereas an infinitive corresponding to a verb selecting ‘be’ does not automatically lead to the use of the auxiliary ‘be’.

*Continuare* seems even more advanced in selecting ‘have’ as the default auxiliary, in that this auxiliary can appear even if it is not motivated by the nature of the subject. This could be an indication of an ongoing grammaticalization auxiliary ‘have’, which is in a more advanced stage in other Romance languages.

### 3.9.4 *Finire*

In the ItTenTen (16) corpus, the CQL query for *finire* ‘finish’ confirms the trend observed with other verbs: ‘have’ is preferred with 2,853 occurrences, but with lower percentages compared to the other semi-auxiliaries analyzed: 91% with ‘have’ and 9% with ‘be’, with 293 occurrences.

*Finire* is unique among semi-auxiliaries of aspect as it combines with various prepositions (such as *a*, *di*, and *per*) and can be found with ‘be’ even when paired with transitive and ‘have’ selecting intransitive infinitives.

***Auxiliary selection properties of the infinitive verb*** The influence of the infinitive verb on auxiliary selection appears to be less pronounced compared to other verbs analyzed thus far, such as *iniziare*, *cominciare*, and *continuare*. ‘Have’ remains the preferred auxiliary when the infinitive is a ‘have’ selecting verb, while ‘be’ is preferred with ‘be’ selecting infinitive verbs.

However, the distribution of ‘have’ and ‘be’ is relatively balanced across almost all configurations depicted in Figure 3.43 and Table 3.62.

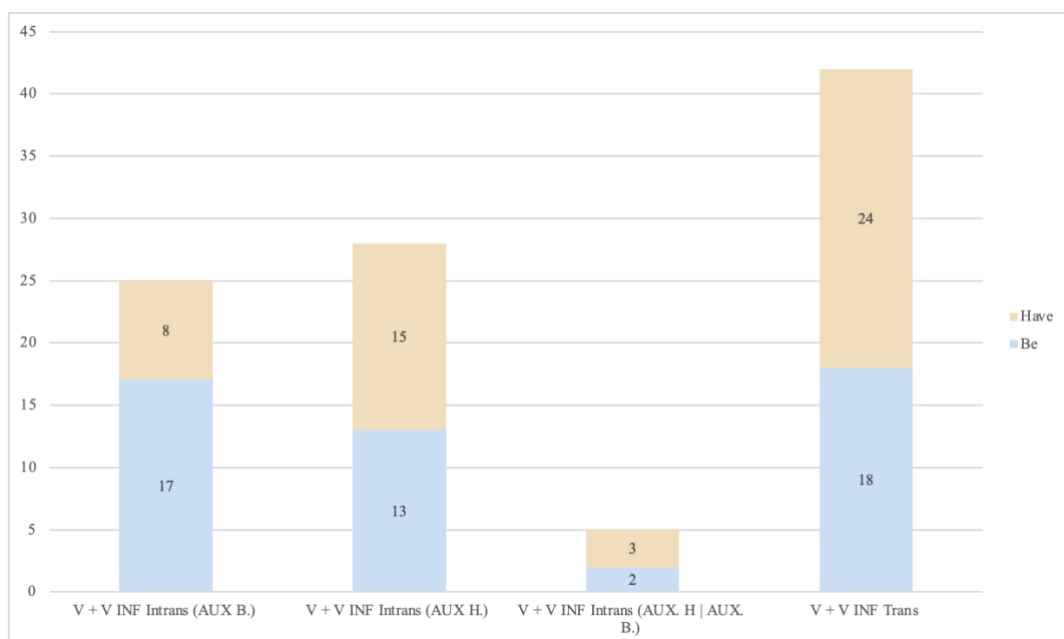


FIGURE 3.43 – Correlation of ‘have’ and ‘be’ with the different types of infinitives.

Factors	Data 'Be'	Data 'Have'	Total
V + V INF Intrans (AUX B.)	17	8	25
V + V INF Intrans (AUX H.)	13	15	28
V + V INF Intrans (AUX H   AUX B)	2	3	5
V + V INF Trans	18	24	42
<b>Total</b>	<b>50</b>	<b>50</b>	<b>100</b>

TABLE 3.62 – Frequency of 'have' and 'be' based on the different types of intransitives.

Figure 3.43 and Table 3.62 show a significant proportion of both 'have' and 'be' auxiliaries irrespective of the auxiliary selection properties of the infinitive verb. The preposition introducing the infinitive seems to have an impact (Figure 3.44).

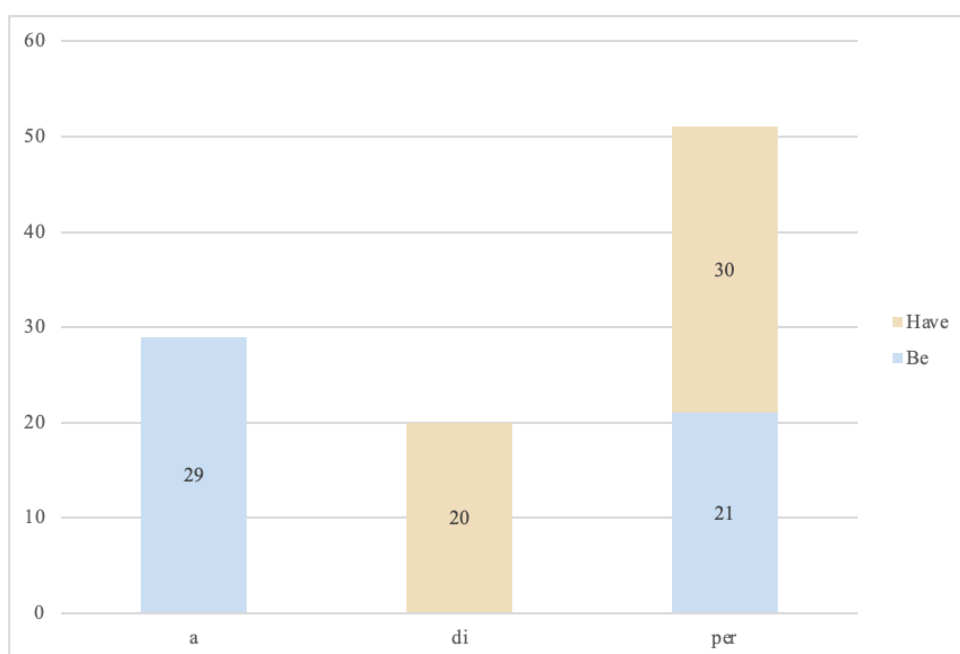


FIGURE 3.44 – Correlation of 'have' and 'be' with the different prepositions heading the infinitive.

The construction *a+infinitive* (i) exclusively features the auxiliary 'be', while the construction *di+infinitive* (ii) exclusively features the auxiliary 'have'. The most notable variation, although with a preference for 'have', is observed in the *per+infinitive* (iii) construction.

#### (I) *FINIRE A + INFINITIVE*

The construction *finire a* exclusively features the auxiliary 'be', when the infinitive verb is transitive (185), as well as with 'have' selecting intransitive verbs (186) and verbs that can

select both auxiliaries as infinitives (187). There are no occurrences of the preposition *a* with intransitive ‘be’ selecting verbs.

- (185) [...] molti studenti sono finiti a studiare ciò che  
 many students be.PRS.3PL finish.PTCP.M.PL to study.INF what that  
 studiano condizionati dagli amici [...].  
 study.PRS.3PL condition.PTCP.M.PL by friends  
 ‘[...] many students ended up studying what they study conditioned by friends [...].’

- (186) [...] ho cercato in tutti i modi di fare qualcosa ma  
 have.PRS.1SG try.PTCP in every DEF.ART ways to do.INF something but  
 alla fine siamo finiti a litigare.  
 finally be.PRS.1PL finish.PTCP.M.PL to fight.INF  
 ‘[...] I tried very hard to do something, but we ended up fighting.’

- (187) [...]i due piccioncini sono finiti a vivere insieme [...].  
 DEF.ART two lovebirds be.PRS.3PL finish.PTCP.M.PL to live.INF together  
 ‘[...] the two lovebirds ended up living together [...].’

*Finire a* in (185), (186), and (187) conveys the meaning of ‘end up doing something’. Regardless of the type of infinitive, this construction appears to emphasize the lack of control over the process. Hence, the only auxiliary found in this context is ‘be’.

## (II) *FINIRE DI* + INFINITIVE

When *finire* combines with the preposition *di*, the only possible auxiliary is ‘have’. Consider (188) as an example of a transitive verb, (189) featuring a ‘have’ selecting intransitive verb, and (190) containing a ‘be’ selecting intransitive verb.

- (188) [...] in un paio di giorni ho finito di leggerlo.  
 in INDF.ART couple of days have.PRS.1SG finish.PTCP to read.INF.it.  
 ‘[...] in a couple of days I finished reading it.’

- (189) Io ho finito di lavorare il 31 [...].  
 I have.PRS.1SG finish.PTCP to work.INF DEF.ART 31  
 ‘I finished working on the 31st [...].’

- (190) L'attore ha finito di truccarsi.  
 DEF.ART.actor have.PRS.3SG finish.PTCP to apply.makeup.REFL.INF

'The actor finished applying makeup.'

The construction *finire di* means 'finish doing something', and hence, implies a kind of control over the action: in the 19 instances of this construction, the subject is either human or non-human but corresponding to an internal cause as responsible for the action. This explains why 'have' is the only accepted auxiliary with these occurrences.

### (III) *FINIRE PER* + INFINITIVE

When the verb *finire* combines with the preposition 'per', the preferred auxiliary changes is conditioned by the auxiliary selection properties of the infinitive verb: 'have' is the preferred auxiliary with transitive verbs like in (191) and intransitive 'have' selecting verbs like in (192), while 'be' is the preferred auxiliary with 'be' selecting ones (193).

- (191) [...] troppa gente ha finito per votare sé stessa [...].  
 too.much people have.PRS.3SG finish.PTCP to vote.INF themselves

'[...] too many people ended up voting themselves [...].'

- (192) Temeva forse che il governo avrebbe finito per  
 fear.IPFV.3SG perhaps that DEF.ART government have.COND.3SG finish.PTCP to

agire diversamente [...].  
 act.INF differently

'He feared perhaps that the government would end up acting differently [...].'

- (193) L'aspetto politico di ognuno di quei problemi ha finito  
 DEF.ART.aspect political of each of those problems have.PRS.3SG finish.PTCP

per diventare non solo preminente, ma totalizzante [...].  
 to become.INF not only preeminent but totalizing

'[...] the political aspect of each of those problems ended up becoming not only preeminent, but totalizing [...].'

In the following section, the possible impact of human nature of the subject will be investigated. The focus will be on the *per*+infinitive construction, as the other constructions have been observed to consistently present only one auxiliary, thus rendering the feature 'human' irrelevant.

**Human and non-human subjects** In contrast to what has been observed for other semi-auxiliaries of aspect, the  $\pm$  HUMAN feature has a less straightforward impact on the choice of the auxiliary used for the perfect tense of *finire per*. Both human and non-human subjects select ‘be’ when the infinitive is ‘be’ selecting verb and ‘have’ when the infinitive is a ‘have’ selecting verb, as depicted in Figure 3.45 and Table 3.63.

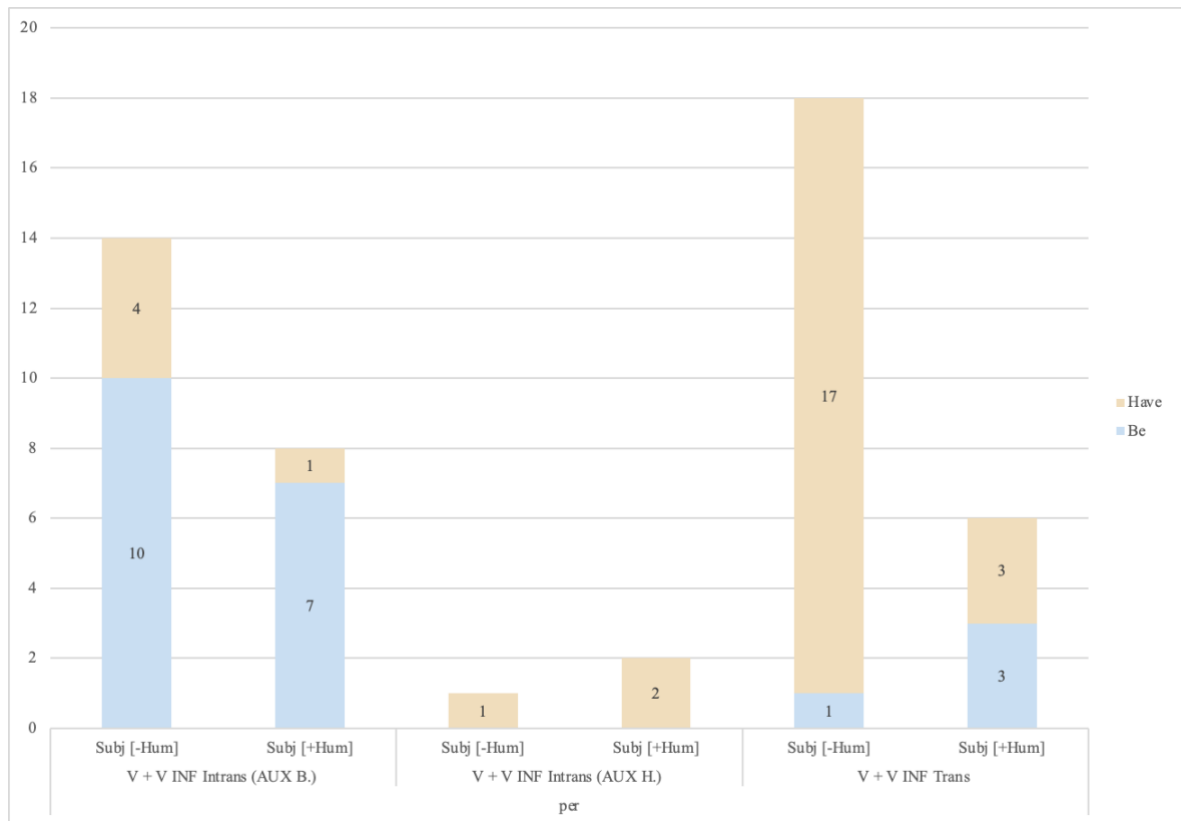


FIGURE 3.45 – Correlation of ‘have’ and ‘be’ with the *per*+infinitive construction, the different types of infinitives and human/non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
[per] [V + V INF Intrans (AUX B)] Subj [-Hum]	10	4	14
[per] [V + V INF Intrans (AUX B)] Subj [+Hum]	7	1	7
[per] [V + V INF Intrans (AUX H)] Subj [-Hum]	/	1	1
[per] [V + V INF Intrans (AUX H)] Subj [+Hum]	/	2	2
[per] [V + V INF Trans] Subj [-Hum]	1	17	18
[per] [V + V INF Trans] Subj [+Hum]	3	3	6
<b>Total</b>	21	28	49

TABLE 3.63 – Frequency of ‘have’ and ‘be’ with the *per*+infinitive construction, based on the auxiliary selection properties of the infinitive verb and the distinction human/non-human subjects.

Figure 3.45 and Table 3.63 show the auxiliary selection based on the nature of the subject. This trend highlights a distinct behavior of *finire per* compared to other semi-auxiliaries.

The following examples show that when the infinitive is a transitive verb, invariably ‘have’ selecting, non-human subjects may cooccur with ‘have’ (194), while human subjects may cooccur not only with ‘have’ (195) but also with ‘be’ (196):

- (194) [...]tutte queste restrizioni della circostanza avevano  
 all these restrictions of. DEF.ART circumstance have.IPFV.3PL  
 finito per irritarlo.  
 finish.PTCP to irritate.INF.him

‘[...] all these restrictions of circumstance had ended up irritating him.’

- (195) [...] io ho finito per darla per scontata.  
 I have.PRS.1SG finish.PTCP to take.INF.it for granted

‘[...] I ended up taking it for granted.’

- (196) La maggior parte delle aziende [...] è finita per  
 DEF.ART most part of. DEF.ART companies be.PRS.3SG finish.PTCP.F.SG to  
 incrementare la popolazione del cimitero degli zombie [...].  
 increase.INF DEF.ART population of. DEF.ART cemetery of.DEF.ART zombie

‘Most companies [...] ended up increasing the population of the zombie cemetery [...].’

(195) and (196) show that there is no difference in meaning whether the auxiliary is ‘have’ or ‘be’.

With ‘be’ selecting verbs, both auxiliaries are possible, but both with human (197) and non-human subjects (198), ‘be’ is more frequent, although ‘have’ also occurs (200).

- (197) I preti sono finiti per essere dei funzionari  
 DEF.ART priests be.PRS.3PL finish.PTCP.M.PL to be.INF of.DEF.ART officials  
 del culto [...].  
 of. DEF.ART cult

‘Priests ended up being cult officials [...].’

- (198) [...] questo mio commento è finito per diventare  
 this my comment be.PRS.3SG finish.PTCP to become.INF

lunghissimo!  
 very long

‘[...] this comment of mine ended up being very long!’

**Agentivity, internal cause and lack of agentivity** Figure 3.46 and Table 5.64, restricted to ‘be’ selecting infinitive verbs, show that proper agentive subjects as well as internal cause subjects are very rare in the construction *finire per* followed by a ‘be’ selecting infinitive verb. As to non-agentive subjects, they tend to cooccur with ‘be’, though ‘have’ is not excluded, indicating the ongoing grammaticalization of ‘have’ as a default auxiliary.

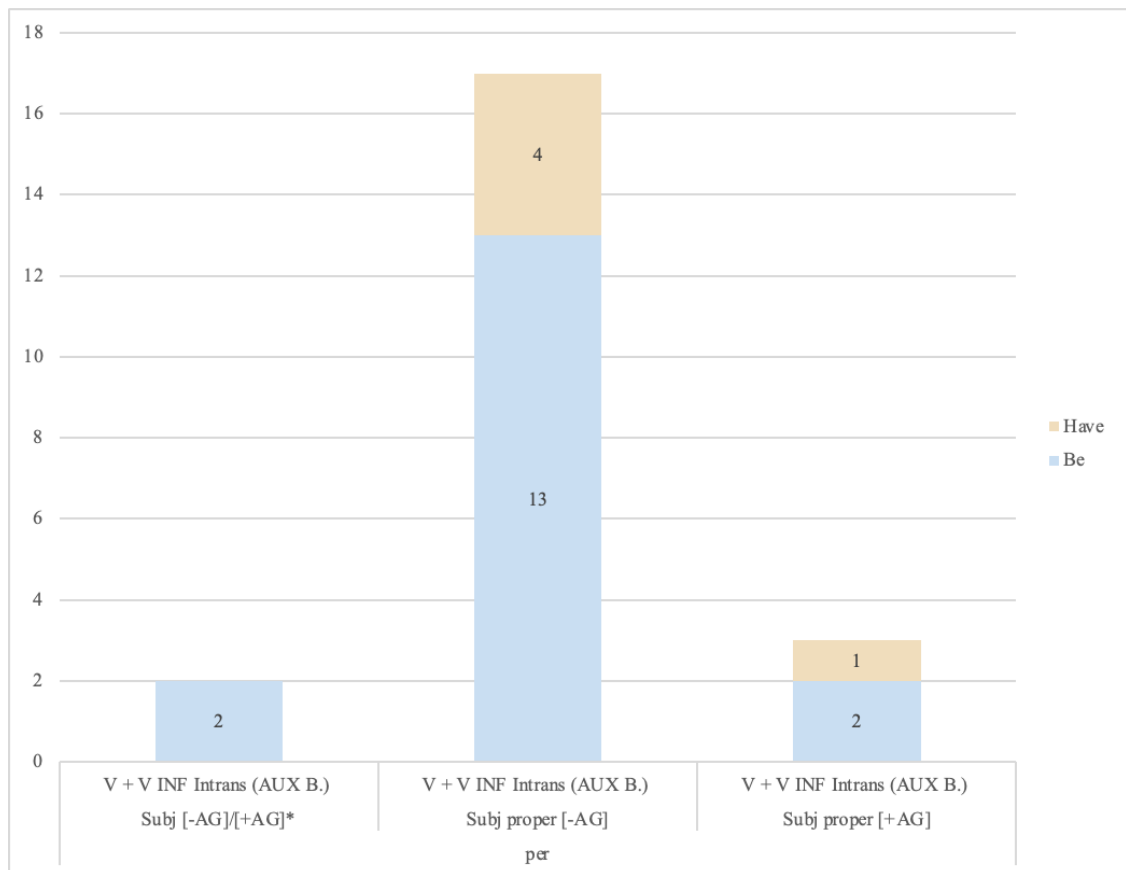


FIGURE 3.46 – Correlation of ‘have’ and ‘be’ with the *per*+infinitive construction, the ‘be-selecting infinitive and agentive/internal case/non-agentive subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
[per] [V + V INF Intrans (AUX B)] Subj [-AG]/[+AG]*	2	/	2
[per] [V + V INF Intrans (AUX B)] Subj [proper [-AG]	13	4	17
[per] [V + V INF Intrans (AUX B)] Subj [proper [+AG]	2	1	3
<b>Total</b>	17	5	22

TABLE 3.64 – Number of occurrences of ‘have’ and ‘be’ with the *per*+infinitive construction, based on the ‘be-selecting infinitive and agentive/internal case/non-agentive subjects.

Consider (200) as an example of a non-agentive subject with ‘have’ and (201) with ‘be’:

(200) [...] il rapporto tra i sessi ha finito per  
 DEF.ART relationship between DEF.ART sexes have.PRS.3SG finish.PTCP to  
 restare impigliato nei suoi risvolti ‘pubblici’.  
 stay.INF entangle.PTCP in-DEF.ART its implications public  
 ‘[...] the relationship between the sexes ended up getting entangled in its “public” implications.’

(201) Il mio romanzo è finito per essere molto più  
 DEF.ART my novel be.PRS.3SG finish.PTCP to be. INF much more  
 rosa di quel che avrei voluto.  
 pink than that which have. COND.1SG want.PTCP  
 ‘My novel ended up being much pinker than I would have liked.’

To conclude, *finire* ‘finish’ is different from the other semi-auxiliaries of aspect. Based on the preposition it is built with, it creates different constructions. The construction *finire di* + infinitive only presents ‘have’ as an auxiliary. It conveys a sort of control over the process (corroborated by the fact that the subjects are all agentive or internal causes).

*Finire a* + infinitive, it only presents ‘be,’ and the construction rather emphasizes a lack of control. The context does not exclude agentive or internal cause subjects.

Finally, the construction *finire per* + infinitive allows variation, especially with ‘be’ selecting verbs. Some instances seem to suggest an ongoing grammaticalization of ‘have’, giving it the status of the default auxiliary. Italian thus appears to be initiating the gradual loss of auxiliary alternation in the infinitive constructions with semi-auxiliaries of aspect, an evolution that is complete in French. This evolution has also been detected for other semi-auxiliaries of aspect.

The human and non-human nature of the subject, as well as the agentivity parameter, do not influence the choice of auxiliary.

Table 3.65 provides a summary of the primary factors influencing auxiliary selection with *finire*.

Infinitive	Agentivity	Human Trait	Auxiliary
<i>Di</i> + infinitive	Agentive, Internal Cause, Non-agentive	Human, Non-human	Have
<i>A</i> + infinitive	Agentive, Internal Cause, Non-agentive	Human, Non-human	Be
<i>Per</i> + infinitive ‘have’	Agentive, Internal Cause, Non-agentive	Human, Non-human	Have
<i>Per</i> + infinitive ‘be’	Agentive, Internal Cause, Non-agentive	Human, Non-human	Be

TABLE 3.65 – Primary factors affecting the selection of the auxiliary in *finire* infinitive constructions

*Finire* indeed stands out as a unique semi-auxiliary, characterized by the presence of multiple constructions based on the preposition. Generally, factors relevant for other verbs, such as the auxiliary selection properties of the infinitive verb, human trait, and agentivity, are not significant with *finire*. The only instance where the infinitive plays a role is the ‘*finire per* + infinitive’ construction.

However, based on the results, the hypothesis is that ‘have’ can be used as a default auxiliary with ‘be’ selecting verbs and non-human and non-agentive subjects, hinting at an ongoing grammaticalization process, even if the occurrences are not sufficient to make a definitive statement. As such, the ‘*finire per* + infinitive’ construction behaves in part similarly to the other ‘semi-auxiliary + infinitive’ constructions analyzed in this study.

### 3.10 INTERMEDIARY RESULTS: THE SEMI-AUXILIARIES OF ASPECT IN INFINITIVE

The auxiliary selection properties of verbs indicating a stage of a process have been analyzed in section 5.7. However parameters related to the nature of the subject, which account for the selection of the auxiliary when these verbs are used as main verbs, proved to be less effective to account for auxiliary selection when these same verbs combine with an infinitive. Hence, it was hypothesized that these verbs function as semi-auxiliaries and that in this status, they may inherit properties from the infinitive acting as the main verb, including its auxiliary selection properties.

A detailed analysis of the different semi-auxiliaries of aspect demonstrated that this hypothesis was corroborated for only some of the considered verbs. As to *finire*, my data show that the auxiliary selection properties of the infinitive verb do not have any impact for the constructions *finire a* + infinitive, selecting only ‘be’, and *finire di* + infinitive, selecting only ‘have’, but do play a role for *finire per* + infinitive.

Secondly, for all other infinitive constructions with an aspectual semi-auxiliary, there is indeed an impact of the auxiliary selection properties of the infinitive on the choice of auxiliary for the semi-auxiliary. But a dissymmetry emerges between ‘have’ and ‘be’: if the infinitive is a ‘have’ selecting verb, the chosen auxiliary will be ‘have’, with a few exceptions. If the infinitive is a ‘be’ selecting verb, both auxiliaries occur.

Thirdly, in this case, the features defining the nature of the subject, such as humanness, agency, and internal cause, come into play, and may still lead to ‘have’.

Finally, ‘have’ may appear as the auxiliary, even when none of the aforementioned factors motivate it. This is especially the case with the semi-auxiliary *continuare*. It thus appears that the use of ‘have’ is the dominant feature in infinitive constructions, while ‘be’ is recessive, leading us to put forward the hypothesis that a grammaticalization of ‘have’ is incipient, evolving to its use as the ‘default’ auxiliary.

As outlined in 3.9, this study does not discount the significance of the main verbs emphasized in works dedicated to restructuring phenomena. *Continuare*, *cominciare*, and *iniziare*, as well as the construction *finire di*, are more associated with the subject’s control over action. Furthermore, recurrent tendencies have been identified, indicating systematic patterns. Firstly, auxiliary selection depends on the infinitive following the past participle: for three of the analyzed verbs (*continuare*, *iniziare*, and *cominciare*), when the infinitive is a ‘have’ selecting verb, whether transitive or intransitive, the auxiliary used is ‘have’.

Another factor is the human trait or agentivity of the subject, at work when the infinitive is a ‘be’ selecting verb: the more human and agentive the subject, the more likely the auxiliary chosen is ‘have’, or there is a more balanced distribution between ‘have’ and ‘be’. However, when the subject is non-human, the preferred auxiliary is typically ‘be’.

Table 3.66 summarizes the most important factors of auxiliary selection with semi-auxiliary verbs.

General Factors	Type of Factor	Different types of verbs
<b>1: INFINITIVE TYPE</b>	<i>'Have' or 'be' selecting infinitives</i>	'Have' selecting (transitives and intransitives): 'have' 'Be' selecting intransitives: 'be'
<b>2: HUMAN TRAIT</b>	<i>Animacy and human trait</i>	[+Hum] with 'have' selecting verbs: 'have' [-Hum] with 'have' selecting verbs: 'have'
		[+Hum] with 'be' selecting verbs: ('have')/'be' [-Hum] with 'be' selecting verbs: 'be'
<b>3: AGENTIVITY</b>	<i>Agentivity, internal cause and lack of agentivity</i>	Agentive, internal cause and lack of agentive subjects with 'have' selecting verbs: 'have' Agentive, internal cause with 'be' selecting verbs: 'have'/'be' Non-agentive subjects with 'be' selecting verbs: 'be'

TABLE 3.66 – Major factors affecting the auxiliary selection for semi-auxiliary verbs (the distribution of 'have' or 'be' is often the result of co-occurrence of more factors).

As depicted in Table 3.66, the most significant influence appears to be exerted by the infinitive: the auxiliary selection properties of infinitive verb determine the selection of the auxiliary. This influence is particularly notable when the verb at the infinitive is a 'have' selecting verb.

Other co-occurring factors come into play, such as the human trait and agentivity, which are contingent upon the type of infinitive.

Another noteworthy aspect that emerged in this study is what has been referred to as the 'grammaticalization of 'have' based on the phenomenon already seen for other languages (Aranovich 2003). This phenomenon is evident with 'be' selecting verbs in all four verbs analyzed. Indeed, many non-agentive non-human subjects can select 'have' with 'be' selecting verbs without a difference in meaning, as it has become a default auxiliary. This confirms a tendency already observed in other Romance languages (Rea 2018; Aranovich 2003).

### 3.11 CONCLUDING REMARKS ON AUXILIARY SELECTION IN ITALIAN

A major distinction that emerged during the analysis of the empirical data was between main verbs and verbs combining with a non-finite verb form such as the infinitive and expressing an

aspectual value. The latter configuration, which has not been considered in the literature on auxiliary selection *per se*, stands out in that the choice of auxiliary is partly determined by the auxiliary selection properties of the infinitive verb, although not entirely.

This is why verbs like *iniziare* integrated into an infinitive construction deserve their name of semi-auxiliaries, since they formally convey some features of the infinitive verb, which acts as the main verb. One of the contributions of my thesis is to have isolated this configuration, in which auxiliary selection is governed by a specific parameter, namely the auxiliary selection properties.

Apart from this distinction, both full verbs and semi-auxiliary verbs can be further identified based on whether they select ‘have’ or ‘be’ as their auxiliary.

Concerning full verbs, they can be distinguished based on the most frequently selected auxiliary: verbs with a prevalence of ‘have’ (e.g., *contare*), verbs with a prevalence of ‘be’ (e.g., *cambiare*), and verbs presenting both with no emerging prevalence (e.g., *suonare*). For full lexical verbs, the semantic parameters of agentivity, internal cause, and human trait influence the choice of the auxiliary. In fact, the internal cause, in combination with the [+HUMAN] feature, can also characterize the subjects when these verbs exhibit an unexpected auxiliary choice (for example, using ‘have’ instead of ‘be’ with *cambiare*).

Semi-auxiliary verbs, instead, are distinguished as ‘have’ and ‘be’ selecting verbs based on the auxiliary that the infinitive selects. For these verbs, the major parameter influencing the selection of the auxiliary is the infinitive: semi-auxiliaries with ‘have’-selecting verbs as infinitives almost exclusively select ‘have’. However, the semantic parameters consisting of agentivity, internal cause, and human trait intervene at a lower hierarchical level: when the infinitive is a ‘be’-selecting verb, the presence of these semantic parameters in the subjects can impact the choice of ‘have’.

In both types of verbs, the human trait and agentivity of the subjects play a significant role: human subjects (highlighted in red in Table 3.67) correlate with ‘have’, showing similarities across all verbs.

Conversely, non-human subjects and the lack of agentivity correlate with ‘be’.

	'HAVE' selection verbs	
	(Full) verbs	Semi-auxiliary verbs
Human trait	[+Hum]: Have	[+Hum]: Have
	[-Hum]: Have/Be	[-Hum]: Have/Be
Agentivity	[+AG]: Have	[+AG]: Have
	[+Internal Cause]: Have/Be	[+Internal Cause]: Have
	[-AG]: Be	[-AG]: Have
	'BE' selection verbs	
	(Full) verbs	Semi-auxiliary verbs
Human trait	[+Hum]: Have	[+Hum]: (Have) Be
	[-Hum]: Be	[-Hum]: Be
Agentivity	[+AG]: Have	[+AG]: Have/Be
	[+Internal Cause]: Have/Be	[+Internal Cause]: Have/Be
	[-AG]: Be	[-AG]: Be

TABLE 3.67 – Factors affecting the auxiliary selection of full and semi-auxiliary verbs.

The [+HUMAN] and [+AGENTIVE] attributes highly correlate with the selection of 'have' as the auxiliary, carrying significant weight. The only relative exception occurs with semi-auxiliaries, where the primary determining factor is the type of the infinitive.

Even within the context of semi-auxiliaries, the nature of the subject still has an indirect influence. With transitive infinitives, the subject tends to be agentive or an internal cause. Similarly, with intransitive infinitives that select 'have', the subject often exhibits the same characteristics. Conversely, with intransitive infinitives that select 'be', the subject tends to be a patient and not agentive (even if human, as in the case of the verb *morire* 'die').

This correlation is less strict because the subject of a transitive verb can also be non-agentive, as illustrated by the verb *subire* 'suffer' in the sentence *Piero ha subito degli insulti* 'Piero suffered insults'.

In addition to the observed tendencies and factors correlating with the selection of 'have' and 'be', one of the most intriguing findings pertains to what determines the distribution of these auxiliaries in peripheral verbs: internal cause subject, which influences both 'be' and 'have' selection full verbs. While internal cause can also have an impact on semi-auxiliary verbs, the most influential factor for this type of verb is the type of infinitive.

In conclusion, the human trait and agentivity emerge as two significant factors influencing auxiliary selection for Italian verbs. Other factors, such as aspect for certain types of verbs (e.g., *suonare*), also play a role, but on a broader scale, the semantic factors of human trait, agentivity, and internal cause are revealed to be crucial.

It is not surprising that, at this stage of the analysis, the distinction between auxiliaries operates at a semantic level: prior to reaching peripheral verbs, the initial syntactic distinction has already been made. As differences become subtle and difficult to discern, semantics assumes a prominent role.

This is why internal cause can account for the usage of both auxiliaries in the majority of contexts: its interpretation, situated between lack of agentivity and agentivity, by virtue of its causative nature, allows for the possibility of leading to the selection of auxiliaries based on the interpretation made.

## CHAPTER 4: CORPUS ANALYSIS OF FRENCH VERBS

This chapter focuses on presenting and discussing the results of corpus analyses conducted to explore variations in the choice between ‘have’ and ‘be’ among French verbs. The selection of verbs followed a two-step process. Initially, a specific Corpus Query Language (CQL) was employed to identify verbs associated with both ‘have’ and ‘be’.

The first CQL used to find verbs associated with both ‘have’ and ‘be’ was: [tag= “N.|P.”] [lemma= “avoir|être”] [tag= “V.P.”]. With this CQL, among the initial 2000 occurrences, I selected the verbs that regularly allow auxiliary alternation. These ten verbs: *changer*, *passer*, *monter*, *diminuer*, *paraître*, *apparaître*, *mûrir*, *baisser*, *descendre*, and *exploser*.

Distinct CQL queries were subsequently applied to each verb. For instance, let’s consider the verb *changer* ‘change’:

(1) [tag= “N.|P.”] [lemma= “avoir”] [word= “changé|changée|changés |changées”]

(2) [tag= “N.|P.\*”] [lemma= “être”] [word= “changé|changée|changés |changées”].

The decision was made to employ Sketch Engine’s ‘random sample’ function for the selection of 50 sentences featuring ‘have’ and 50 sentences with ‘be’, whenever possible. This aimed to obtain a limited yet representative number of occurrences from the corpus. Following the methodology applied in Italian, the context filter ‘no noun in the first five words on the right’ was utilized to minimize transitive contexts. In cases where the initial 50 sentences yielded insufficient data meeting the analysis criteria, the random sample size was expanded. The ‘random sample’ function may not be selected in specific instances with ‘be’ when its application fails to yield sufficient data, given the anticipated lower frequency of ‘be’ as an auxiliary in French.

The ten identified verbs, namely *changer*, *passer*, *monter*, *diminuer*, *paraître*, *apparaître*, *mûrir*, *baisser*, *descendre*, and *exploser*, can be categorized into three semantic classes<sup>19</sup> regrouping verbs sharing a similar distribution of the auxiliaries ‘have’ and ‘be’:

---

<sup>19</sup> The terminology has been based on Levin’s (1993) classification of verbs.

- (i) Spatial Motion verbs (Levin 1993 : 263) : *passer, monter, descendre, baisser*
- (ii) Change-of-state verbs (Levin 1993: 245): *changer, mûrir, diminuer, exploser*
- (iii) Verbs of Appearance (Levin 1993: 258), having a copulative construction: *paraître, apparaître*

The third class of verbs raises problems insofar as it comes in two clearly distinct constructions, the second being associated with an adjectival predicate.

With respect to the two first classes, at the exception of *exploser*, all verbs are so-called ‘labile verbs’, showing a causative alternation: that is, they can be used both transitively and intransitively, with the requirement that the direct object of their transitive use corresponds to the subject of their intransitive use:

- (3) a. *Les banques **ont monté/diminué** les taux d'intérêt/*
- (4) b. *Les taux d'intérêt **ont monté/diminué**.*
- (5) a. *Le soleil **a mûri** le raisin.*
- b. *Le raisin **a mûri**.*

As transitive verbs invariably select ‘have’, my analysis will be restricted to intransitive constructions of these verbs, for which auxiliary alternation is possible.

In line with the above observations for Italian, it might be expected that the crucial factor is the more or less causative nature of the subject: if the subject is agentive or internal cause, ‘have’ should be preferred, whereas ‘be’ is selected in the absence of agentivity and causativity.

- (6) a. *La pomme **a mûri**.*
- b. *Ce projet **est mûri**.*

The French data show, however, a more complex picture. In the first place, it is unclear whether spatial motion verbs and change-of-state verbs behave in the same way with regard to auxiliary selection: the former class tends to favor ‘be’ more than the latter.

Secondly, the relationship with the transitive construction may be more or less close. Some verbs, such as *mûrir*, are infrequent in the transitive construction. Other intransitive verbs of the list of selected verbs have developed polysemy and show meanings that do not exist for the transitive verb.

And last but not least, the use of the auxiliary ‘have’ is more grammaticalized in French than in Italian and, as a result, the relationship between ‘be’ and the feature of unaccusativity is less

straightforward. A detailed corpus study of the auxiliary selection of these French verbs is therefore necessary and will be the subject of this chapter.

**Sample size and representativeness** Prior to conducting a detailed analysis of each verb, it is essential to consider the auxiliaries associated with different verbs. I examined 50 sentences with ‘have’ and 50 with ‘be’ to identify the most relevant factors. It is important to note that these verbs may exhibit varying percentages of occurrences with ‘have’ and ‘be’ in the corpus. To ensure the representativeness of the corpus, I therefore assessed whether the distribution of transitivity and intransitivity in sentences containing verbs that allow for auxiliary selection significantly differed, using contingency tables, considered to display the frequency distribution of categorical variables.

Initially, the first 100 occurrences were retrieved using a specific command in the Sketch Engine CQL query. Two features—type of auxiliary and (in)transitivity or passive construction—were manually annotated in a spreadsheet. The chi-square test was then computed using JASP<sup>20</sup>, an open-source statistical program developed by the University of Amsterdam.

Among the ten analyzed verbs, namely *changer* ‘change’, *passer* ‘go on’, *monter* ‘go up’ ‘increase’, *diminuer* ‘go down’ ‘decrease’, *paraître* ‘seem’ ‘appear’, *apparaître* ‘seem’ ‘appear’, *mûrir* ‘mature’, *baisser* ‘go down’ ‘decrease’, *descendre* ‘go down’, and *exploser* ‘explode’, two distinct situations have been identified: verbs for which the difference in auxiliary selection is statistically significant with respect to the feature of (in)transitivity, and others for which it is not.

The two hypotheses are as follows: the null hypothesis (H0) posits that there is no association between the verbal construction (intransitive vs. transitive) and the choice of the auxiliary. The alternative hypothesis (H1) suggests that an association between the two does exist.

Let us proceed by examining the first case, starting with the verb *changer* ‘change’.

<b>Contingency Tables</b>			
	<b>TRANS</b>		
<b>AUX</b>	<b>0=intransitive</b>	<b>1=transitive</b>	<b>Total</b>
0=have	43	48	91
1=be	9	0	9
Total	52	48	100

TABLE 1 – Contingency table for *changer*

<sup>20</sup> JASP Team (2024). JASP (Version 0.18.3)

In the first 100 occurrences of *changer* using the CQL [tag= “N.|P.”] [lemma= “avoir|être”] [word= “changé|changée|changés |changées”], 43 intransitive constructions selecting ‘have’ and 9 intransitive constructions selecting ‘be’ were observed. All 48 transitive constructions opted for ‘have’. The calculated  $p$ -value in this case is 0.003, indicating that  $p$ -value < 0.005.

This significant difference demonstrates that ‘be’ is exclusively used for intransitive constructions, while ‘have’ emerges as the predominant auxiliary for both transitive and intransitive constructions. Therefore, the null hypothesis is rejected, and the alternative hypothesis (H1) is accepted: there is indeed an association between the construction and the auxiliary, where the intransitive constructions select ‘have’ and the transitive constructions also select ‘have’.

Contrastingly, among the verbs with a non-statistically significant  $p$ -value, we find *mûrir*.

Contingency Tables			
	TRANS		
AUX	0=intransitive	1=transitive	Total
0=have	82	14	96
1=be	4	0	4
Total	86	14	100

TABLE 2 – Contingency table for *mûrir*

For *mûrir*, the  $p$ -value of 0.410 suggests that the observed difference is not statistically significant. Therefore, the null hypothesis (H0) has to be accepted, indicating that the association between the construction and the choice of the auxiliary is not clearly evident. In fact, since most sentences are intransitive (86 vs. 14 transitive), it is challenging to discern a clear distribution of auxiliaries based on the construction. The auxiliary preferred by the intransitive is ‘have’, while for the transitive the only possible auxiliary is ‘have’.

The presented tables underscore the predominance of ‘have’ in French, although ‘be’ also occurs, as my qualitative analysis will elaborate. Among the ten mentioned verbs (*changer*, *passer*, *monter*, *diminuer*, *paraître*, *apparaître*, *mûrir*, *baisser*, *descendre*, and *exploser*), a statistically significant association is observed with a significant  $p$ -value.

In cases where the  $p$ -value is insignificant, it may suggest a need for additional data to accurately assess representativeness. Certain verbs like *paraître* and *apparaître* are either non-analyzable or do not exhibit variability in construction or auxiliaries within the first 100 occurrences, hindering comprehensive analysis.

The intricate nature of these data shows the necessity for in-depth analysis in order to understand why a particular auxiliary is selected despite a default one for a given verb.

#### 4.1 SPATIAL MOTION VERBS

Spatial motion verbs, specifically *passer*<sup>21</sup>, *monter*, *baisser*, and *descendre*, appear in different constructions, exhibiting a significant degree of polysemy in this corpus. These verbs are commonly used in both their spatial motion construction and quantitative construction. When employed in their spatial motion construction, and primary sense, i.e., motion in space, they resemble the core verbs in Italian and are expected to favor the auxiliary ‘be’.

##### 4.1.1 *Passer*

In FrTenTen (17), the verb *passer* ‘go on’ is prominently attested with both ‘have’ and ‘be’. After applying the part-of-speech filter, the corpus reveals a 12% occurrence with ‘have’ versus 88% with ‘be’, with 8,444 occurrences of ‘have’ and 62,364 occurrences of ‘be’.

The verb *passer* is used in three distinct constructions, each corresponding to different meanings: spatial motion (i), quantitative evolution (ii), and a temporal meaning (iii). The type of construction dictates the auxiliary selection of *passer*. Figure 4.1 summarizes the trend of auxiliary distribution related to the different constructions.

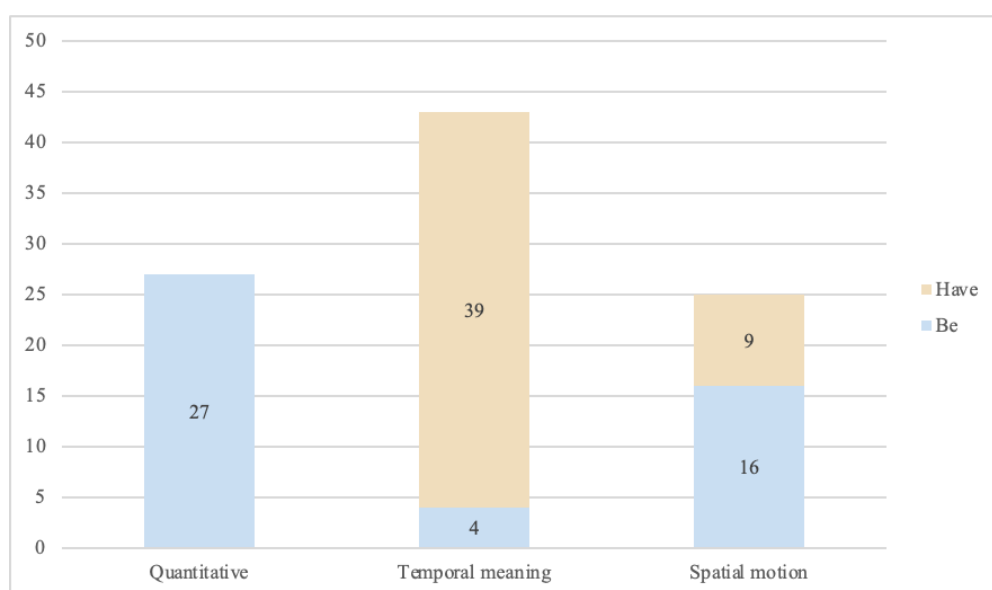


FIGURE 4.1 - Correlation between ‘have’ and ‘be’ and construction type.

<sup>21</sup> *Passer*, *descendre*, and *monter* have been identified as verbs demonstrating alternation, as noted by Heidinger (2015: 280). However, the underlying reasons for this alternation are not explicitly addressed.

Factors	Data 'be'	Data 'have'	Total
Spatial motion	16	9	24
Quantitative evolution	27	/	27
Temporal meaning	4	39	43
<b>Total</b>	<b>47</b>	<b>48</b>	<b>95</b>

TABLE 4.1 – Number of occurrences with ‘have’ and ‘be’ according to construction type.

As Table 4.1 illustrates, there are 95 sentences, while another 5 sentences in the corpus do not fit into either of these categories but represent a general figurative use of the verb *passer* (e.g., *j’ai passé de 13H30 à 00H10 à installer cette nouvelle carte mère* - ‘I spent from 13:30 to 00:10 installing this new motherboard’). Spatial motion constructions prefer ‘be’, even with some uses presenting ‘have’; quantitative evolution constructions prefer ‘be’, and finally, constructions with temporal meaning opt for ‘have’.

(i) **Spatial motion construction** The verb *passer* can be used in spatial motion construction. Specifically, it displays the pattern ‘passer + LOCATIVE ADVERBIAL’ construction, with the auxiliary ‘be’ consistently selected in this context. Consider the following sentences as illustrative examples.

- (1) Je suis passé par là, pourquoi vouloir y retourner?  
 I be.PRS.1SG pass.PTCP over there why want.INF LOC go.back.INF

‘I passed by there, why would I want to go back?’

- (2) [...] c’est pour vérifier qu’un sanglier est passé ?  
 it.be.PRS.3SG to check.INF that.INDF.ART wild boar be.PRS.3SG pass.by.PTCP

‘[...] it’s to check if a wild boar has passed by?’

In (2), the specific locative adverbial is not explicitly specified. However, it can be inferred from the lexical meaning of the subject that the verb is employed within the spatial motion construction. The auxiliary ‘be’ consistently prevails, in spite of sporadic instances of ‘have’, which can be due to the fact that ‘have’ is the default auxiliary in French.

(ii) **Quantitative evolution construction** The quantitative evolution construction prototypically involves the verb *passer* followed by a complement expressing quantity.

However, the quantity can also remain unexpressed. In this context, the exclusive auxiliary selected is ‘be’, as exemplified by instances (3) and (4):

- (3) La part des salaires [...] est passée de 76%  
DEF.ART share of.DEF.ART wages be.PRS.3SG pass.PTCP.F.SG from 76%

à 60% entre 1980 et 1998.  
to 60% between 1980 and 1998.

‘[...] The share of wages [...] decreased from 76% to 60% between 1980 and 1998.’

- (4) En France, les sociétés disposant d’une page Facebook sont  
In France, DEF.ART companies having of.INDF.ART page Facebook be.PRS.3PL

passées de 4,12% à 16% durant l’année 2013.  
pass.PTCP.F.PL from 4,12% to 16% during DEF.ART.year 2013

‘In France, the companies having a Facebook page went from 4,12% to 16% during the year 2013.’

The occurrences in (3) and (4) demonstrate that regardless of the nature of the subject, whether human or non-human, the consistently chosen auxiliary is ‘be’.

**(iii) Temporal meaning** The third construction associated with the verb *passer* conveys a temporal meaning. In contrast to the other two constructions, it predominantly selects ‘have’.

Consider the following sentences.

- (5) Deux semaines ont passé, je ne suis pas revenue.  
two weeks have.PRS.3PL pass.PTCP I not be.PRS.1SG NEG come.back.PTCP.F.SG

‘Two weeks passed and I didn’t come back.’

- (6) Les heures avaient passé [...] Cela commençait à lui peser.  
DEF.ART hours have.IPFV.3PL pass.PTCP That begin.IPFV.3SG to him/her weigh.INF

‘The hours had passed [...]. It was beginning to weigh on him/her.’

In this scenario, ‘be’ is sporadic, whereas ‘have’ is prevalent. Another useful element for understanding the selection of different auxiliaries with the verb *passer* is the nature of the subject.

While the spatial motion construction prefers a human subject (7), although possible with non-human subjects presenting [+HUMAN] attributes (8), non-human subjects are most prevalent in constructions related to quantitative evolution and temporal meanings. In the latter

meaning, the subject corresponds to temporal units of measurement ('months', 'years' etc.). Therefore, the subject serves as a co-occurring factor in determining the construction used.

#### SPATIAL MOTION VERBS | SUBJECT [+HUM]

- (7) Oui, je suis passée par là mais je n'ai réussi à m'identifier à eux.  
 yes I be.PRS.1SG pass.PTCP.F.SG through there but I not.have-PRS.1SG  
 succeed.PTCP to me.identify.INF with them

'Yes, I went through that, but I couldn't identify with them.'

#### SPATIAL MOTION VERBS | SUBJECT [-HUM]

- (8) L'Associated Press précise qu'un autre petit astéroïde  
 DEF.ART.Ass Press point.out.PRS.3SG that.INDF.ART other small asteroid  
 était passé au plus près de notre planète [...].  
 be.IPFV.3SG pass-PST.PTCP at.the.closest of our planet

'The Associated Press specifies that another small asteroid had passed the closest to our planet [...].'

The quantitative evolution construction conveys a derived meaning, as the verb is not used in combination with a locative adverbial but rather a complement specifying quantity. Even in this case, the consistently chosen auxiliary is 'be' (9, 10).

#### QUANTITATIVE EVOLUTION | SUBJ [+HUM]

- (9) Mais également, une tension qui oscille vers le haut  
 But also INDF.ART tension that oscillate.PRS.3SG upwards DEF.ART high  
 de 11.7 je suis passée soudainement à 15.9.  
 from 11.7 I be.PRS.1SG pass.PTCP.F.SG suddenly to 15.9

'But also, a tension that oscillates upwards from 11.7 I suddenly went to 15.9'

#### QUANTITATIVE EVOLUTION | SUBJ [-HUM]

- (10) [...] le taux de chômage pour l'ensemble des actifs  
 DEF.ART rate of unemployment for DEF.ART.whole of DEF.ART workforce  
 est passée de 7,5% à 9,7%.  
 be.PRS.3SG pass.PTCP.F.SG from 7,5% to 9,7%

'[...] the unemployment rate for the workforce as a whole rose from 7.5% to 9,7%.'

As (10) shows, this construction is more likely to have a non-human subject. Human subjects are possible, as in (9), but because the blood pressure values have been attributed to the subject ('I suddenly went to 15.9' would have been 'my blood pressure values went to 15.9'). Nevertheless, this construction, depicting a quantitative evolution, exclusively employs 'be.' In contrast, in sentences with temporal meaning, the meaning is also derived, but it selects 'have' (11), and the subject is invariably non-human.

#### TEMPORAL MEANING | SUBJ [-HUM]

(11=5) Deux semaines ont passé, je ne suis pas  
 two weeks have.PRS.3PL pass.PTCP I not be.PRS.1SG NEG  
 revenue.  
 come.back-PTCP.F.SG

'Two weeks have passed, I have not come back.'

Apart from a few exceptions, temporal meaning constructions predominantly favor the auxiliary 'have.'

In summary, 'have' is favored in the construction involving a derived meaning relating to time, whereas the quantitative evolution constructions and the spatial motion construction of spatial motion prefer 'be'. Some instances of 'have' within a spatial motion construction have been retrieved in the corpus. Example (12) is a sentence from a book published in 1909:

(12) La fille du Kimry a passé [...] Qui résiste  
 DEF.ART daughter of DEF.ART K. have.PRS.3SG pass-PTCP. who resist.PRS.3SG  
 à sa voix?  
 to her voice

'The daughter of the Kymri has passed! Who can resist her voice?'

Given the preferential use of 'be' with *passer* in the spatial motion construction, the use of 'have' with *passer* in the spatial motion construction is too sporadic to warrant analysis.

As anticipated, the constructions analyzed so far are associated with meanings that reflect the possible use of the construction and the types of subjects involved. Based on the construction and associated meaning, the spatial motion construction prefers human subjects, while the other two constructions prefer non-human subjects.

This association between construction, meaning, and subject type has an impact on auxiliary distribution based on the nature of the subject: human subjects would prefer 'be' due to their

spatial motion-primary meaning construction, while non-human subjects will exhibit more variation depending on whether it is a quantitative or temporal construction.

**Human and non-human subjects** The majority of the occurrences of the sample involve non-human subjects, and they exhibit more variation as they appear both in the quantitative (selecting ‘be’) and temporal (selecting ‘have’) construction. In contrast, human subjects tend to prefer ‘be’ (Figure 4.2).

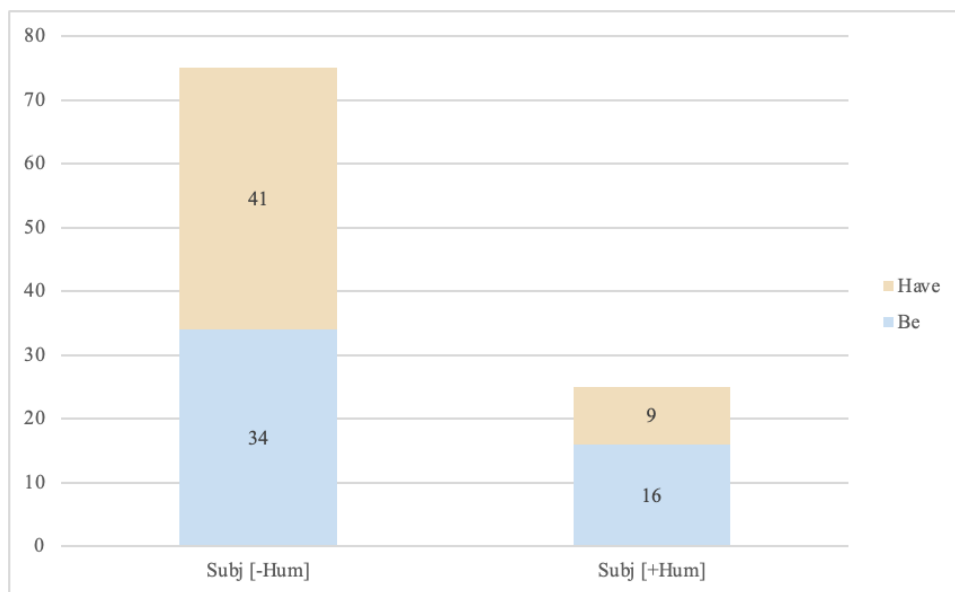


FIGURE 4.2 - Correlation between ‘have’ and ‘be’ and the human or non-human nature of the subject.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	34	41	75
Subj [+Hum]	16	9	25
<b>Total</b>	50	50	100

TABLE 4.2 – Frequency of ‘have’ and ‘be’ according to subject’s nature(human/non-human).

In the dataset featuring non-human subjects, a comparable number of occurrences with both ‘have’ and ‘be’ is observed, albeit with a slight preference for ‘have’. On the contrary, instances involving human subjects exhibit twice as many occurrences with ‘be’ compared to ‘have’. This discrepancy can be elucidated by considering the previously discussed constructions.

The construction with a temporal meaning primarily selects ‘have’ and is more prevalent with non-human subjects. As pointed out before, the construction indicating quantitative evolution exclusively selects ‘be’. Thus, the balanced occurrence of ‘have’ and ‘be’ with non-

human subjects can be attributed to their association with distinct constructions (quantitative in (13) and temporal in (14)):

- (13) En une décennie la part de marché occupée par les  
 In INDF.ART decade DEF.ART part of market occupy.PTCP.F.SG by DEF.ART  
 véhicules essence est passée de 25% à presque 50%.  
 vehicles essence be.PRS.3SG pass.PTCP.F.SG from 25% to almost 50%

‘In the space of a decade, the market share of petrol vehicles went from 25% to almost 50%.’

- (14) Bon d’accord... le temps a passé depuis.  
 Good all right DEF.ART time have.PRS.3SG pass.PTCP since.then

‘Good, all right...time has passed since then’.

When the subjects are human, occurrences are particularly associated with the spatial motion construction or the quantitative evolution. Consider instances (15) and (16), whose analysis is more complex.

- (15) Nous sommes régis par une morale qui nous  
 we be.PRS.1PL govern.PTCP.M.PL by INDF.ART morality that us

rend malheureux![...] comme tout le monde, je  
 make.PRS.3SG unhappy as all DEF.ART people I

suis passée par là.  
 be.PRS.1SG pass. PTCP.F.SG over there

‘We are governed by a morality that makes us unhappy! [...] like everyone else, I’ve been there.’

- (16) La population du Nord-Soudan est passée [...] de  
 DEF.ART population of.DEF.ART North Sudan be.PRS.3SG pass. PTCP.F.SG from

7.48 M à 21,4 M.  
 7,48 M to 21,4 M

‘The population of North Sudan [...] went from 7,48 M to 21,4 M.’

In example (16), a comparable pattern emerges as observed in other instances within the quantitative evolution construction featuring a non-human subject (as *population* refers to a demographic number, not to the people *stricto sensu*). Conversely, example (15) presents a unique and ambiguous case within the spatial construction.

In contrast to the preceding examples characterized by a spatial construction with a primary and literal meaning, this specific spatial construction maintains its primary meaning attributed by the construction (*passer par là*) while incorporating a figurative interpretation. This interpretation is considered primary due to the fact that the verb is embedded in a spatial motion construction, with an explicit locative adverbial.

**Agentivity or lack of agentivity** In the case of *passer*, the selection between ‘have’ and ‘be’ based on agentivity is consistent with other previously analyzed factors, including the categorization into three construction types (spatial motion, quantitative evolution, and temporal meaning), and the associated distinction between primary and derived meanings.

Expectedly, a higher degree of variation is observed with non-agentive subjects, which are predominantly associated with constructions conveying temporal meaning and qualitative evolution. As shown above, ‘be’ is typically favored in the context of quantitative evolution, while ‘have’ is strongly preferred in constructions with temporal meaning.

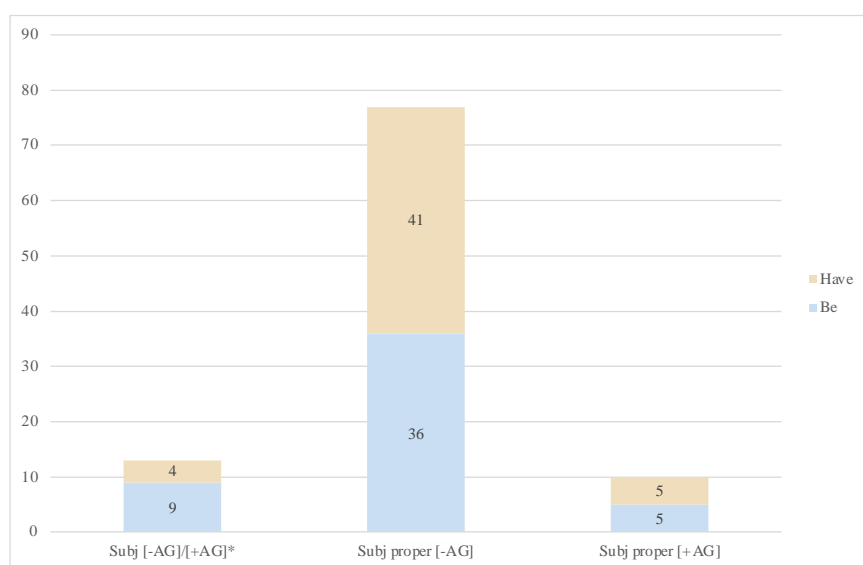


FIGURE 4.3 - Correlation of ‘have’ and ‘be’ in relation to internal cause and agentivity.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG]/ [+AG]* <sup>22</sup>	9	4	13
Subj proper [-AG]	36	41	77
Subj [+AG]	5	5	10
<b>Total</b>	50	50	100

TABLE 4.3 – Total occurrences of ‘have’ and ‘be’ related to internal cause and agentivity.

<sup>22</sup> Subjects characterized by *Internal causation*= Subj [-AG]/[+AG] \* have been distinguished from proper non-agentive subjects.

The table 4.3 serves as an apt illustration of the analysis conducted thus far. Examples (17) and (18) provide instances of ‘have’ (temporal meaning) and ‘be’ (quantitative evolution), respectively:

(17) Cinq ans ont passé [...].  
 five years have.PRS.3PL pass.PTCP

‘Five years have passed [...].’

(18) Ainsi, le taux de remboursement de plus de 450 produits  
 thus DEF.ART rate of reimbursement of more than 450 products  
 est passé de 65% à 35% en avril 2003 [...].  
 be.PRS.3SG pass-PTCP from 65% to 35% in April 2003

‘Thus, the reimbursement rate for over 450 products dropped from 65% to 35% in April 2003 [...].’

In general, instances with a non-agentive subject favoring ‘have’ typically evoke the passage of time, while those with a non-agentive subject favoring ‘be’ typically represent quantitative evolution.

Predominantly, agentive subjects favor spatial motion verb readings, with some exceptions including cases of quantitative evolution with a human subject. In these instances, both ‘have’ and ‘be’ can be utilized, but without having the same frequency. Notably, *passer* as a motion verb tends to select ‘be’, particularly evident in cases involving internal causes.

Examples (19), (20), and (21) exemplify the ‘be’ selection.

(19) Je consomme mes 2 Go tous les mois donc je suis  
 I use.up.PRS.1SG my 2 Go every DEF.ART months so I be.PRS.1SG

passé au 6 Go.  
 pass.PTCP at.DEF.ART 6 Go

‘I use up my 2 GB every month, so I switched to 6 GB’.

(20=7) Oui, je suis passée par là mais je  
 yes I be-PRS.1SG pass-PTCP.F.SG through there but I

n’ai réussi à m’identifier à eux.  
 not.have.PRS.1SG succeed.PTCP to me.identify.INF with them

‘Yes, I went through that, but I couldn’t identify with them.’

- (21) Il est donc important de parler avec ceux qui sont  
 it be.PRS.3SG so important to speak.INF with those who be.PRS.3PL  
 passés par là et de constater qu'ils ne sont pas  
 pass.PTCP.M.PL over there and to see.INF that.they not be.PRS.3PL NEG  
 si malheureux.  
 so unhappy

‘So it is important to speak with those who have been through it and to see that they are not so unhappy.’

In (19) and (20) present instances where ‘be’ is selected with agentive subjects, representing a quantitative evolution construction with human subjects and a spatial motion occurrence, respectively. (19) presents a construction that recalls the spatial motion one (*passer de x à x* ‘move from x to x’), which could explain the use of ‘be’ even if the construction is quantitative.

In (21), the occurrence involves an internal cause subject. Although the verb is used within a spatial motion construction, it conveys a figurative interpretation (*passer par là* in the sense of ‘already experienced a situation’, ‘pass through something’).

Occurrences with spatial motion verbs conveying a figurative meaning are not uncommon, as demonstrated in (22):

- (22) Imaginez maintenant tout ce qui vous attend [...]. Vous êtes  
 imagine.PRS.2PL now all it that you wait.PRS.3SG you be.PRS.2PL  
 passées par le plus obscur, maintenant c'est  
 pass.PTCP.M.PL through DEF.ART more dark now it.be.PRS.3SG  
 la Lumière.  
 DEF.ART Light

‘Imagine now all that awaits you. You have passed through the darkest; now it’s the Light.’

With internal cause subjects, constructions involving the verb *passer* adopt typical movement patterns but with altered semantics. The prevalent interpretation is ‘to come before someone’ or ‘to have prior experience of something.’

An intriguing example with an internal cause subject is offered by (23), featuring a spatial construction with a figurative meaning.

(23)	Il	sera	passé	progressivement,	insidieusement,	de façon	plus	
	He	be.FUT.3SG	pass.PTCP	gradually	insidiously	in way	more	
	ou	moins	rapide	selon	la	drogue	utilisée,	du
	or	less	quick	depending_on	DEF.ART	drug	use PTCP.F.SG	from
	plaisir	à la	souffrance					
	pleasure	at DEF.ART	suffering					

‘He will have transitioned gradually, insidiously, and at varying speeds depending on the drug used, from pleasure to suffering.’

The expression *passer du plaisir à la souffrance* ‘transitioning from pleasure to suffering’ metaphorically encapsulates the concept of ‘moving from one state to another’.

Overall, it appears that the nature of the subject does not impact the choice of the auxiliary *per se*. Instead, it is the three previously identified constructions that constitute the major parameter, with the nature of the subject being a side-effect of the chosen construction.

**Adverbials** Most instances of adverbials belong to the following types: quantity adverbials, temporal adverbials, and time + duration adverbials. However, their presence or absence does not significantly influence auxiliary selection. Nevertheless, exceptions to this trend do exist.

**QUANTITY ADVERBIALS** Among the 25 occurrences with an adverbial of quantity, ‘be’ is predominantly selected. However, the preference for ‘be’ is not directly linked to the adverbial of quantity itself but rather to the construction in which it is embedded. Specifically, it aligns with the quantitative evolution construction, a pattern which features exclusively the auxiliary ‘be’.

**TEMPORAL ADVERBIALS** Based on the data analyzed, it appears that the adverbial of time (for example, *en avril 2003* ‘in April 2003’) does not exert a significant influence on auxiliary selection. With only 11 occurrences out of 100 sentences, the frequency is too low to draw definitive conclusions regarding the adverbial of time’s impact.

**TIME + DURATION ADVERBIALS** Adverbials of time plus duration, such as *de mes 14 ans à maintenant* (‘from my 14 years old to now’, present in the corpus), often convey a resultative meaning by indicating both a specific time period and its duration. However, with the verb *passer*, adverbials of time combined with duration do not appear to significantly impact auxiliary selection.

This conclusion is drawn from the limited number of occurrences, totaling only 11 instances. Within these occurrences, 9 instances favor the auxiliary ‘be.’ However, this bias towards ‘be’ can largely be attributed to the prevalence of quantitative evolution constructions in this sample, a pattern already identified in my analysis.

In summary, the verb *passer* demonstrates a notable sensitivity to the syntactic constructions it operates within. The primary determinant of auxiliary selection is the type of construction employed, i.e., spatial motion, qualitative evolution, or temporal meaning.

Other factors typically considered, such as the human trait or agentivity, emerge as secondary considerations in auxiliary selection, often serving as consequential factors within the broader framework of the verb’s usage patterns.

The primary factors influencing the choice of ‘have’ and ‘be’ are presented in table 4.4.

<b>Construction</b>	<b>Human trait</b>	<b>Agentivity</b>	<b>Auxiliary</b>
Spatial motion	Human	Agentive →	Be
Quantitative evolution	Non-human	Not agentive →	Be
Temporal meaning	Non-human	Not agentive →	Have

TABLE 4.4 – Primary factors influencing the choice of ‘have’ and ‘be’ for *passer*.

Table 4.4 outlines the interaction between two crucial factors: syntactic constructions and semantic factors related to the nature of the subject. Spatial motion verbs, typically associated with a primary meaning, tend to favor ‘be’ as the auxiliary, with occasional exceptions, as previously noted. Human subjects corresponding to agents predominate in this construction.

Quantitative evolution constructions exhibit a correlation with ‘be’ and are associated with derived meanings. The canonical subject in this construction lacks agency and is typically non-human. A second derived construction, conveying a temporal meaning, has equally a subject lacking agency and typically non-human, but selects ‘have’ as the preferred auxiliary.

#### 4.1.2 *Monter*

In the FrTenTen (17) corpus, an analysis of the usage distribution of ‘have’ and ‘be’ with the verb *monter* ‘go up’ shows that this verb allows for the use of both auxiliaries but with a preference for ‘be’: ‘be’ predominates, with 80% occurrences of ‘be’ compared (4,550 occurrences) to 20% with ‘have’ (1,443) after applying the part-of-speech filter. These

statistical data diverge from Italian, where the equivalent verb typically exclusively selects ‘be’ in its intransitive construction.

*Monter* is used primarily in spatial motion and quantitative evolution constructions, as depicted in Figure 4.4.

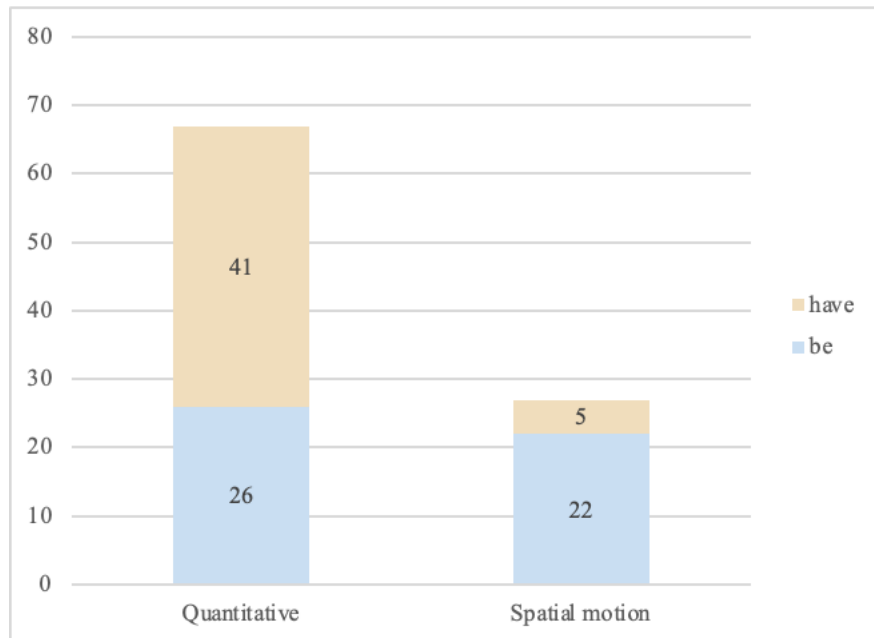


FIGURE 4.4 - Correlation between the construction’s type and the selection of ‘have’ and ‘be’.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Spatial motion</b>	22	5	27
<b>Quantitative evolution</b>	26	41	67
<b>Total</b>	48	46	94

TABLE 4.5 – Total occurrences relative to ‘have’ and ‘be’ based on the construction.

As illustrated in Table 4.5, there are six sentences lacking attributions, as they do not belong to either a spatial motion or a quantitative evolution construction. In the spatial motion construction, the preferred auxiliary is ‘be’, while in the quantitative evolution construction, the favored auxiliary is ‘have’.

***Spatial motion construction*** When the verb is used in its spatial motion construction, the favored auxiliary is ‘be.’ This is evident in examples (24) and (25).

- (24) Un bon souvenir, ce Crêt de Chalam [...] On y était  
 INDF.ART good memory this C. de C. 3SG LOC be.PRS.3SG  
 monté avec Pascale il y a quelques années  
 climb.PTCP with P. 3SG LOC have.PRS.3SG a.few years  
 en hiver.  
 in winter

‘A good memory, this Crêt de Chalam. [...] We had climbed it with Pascale a few years ago, in winter.’

- (25) La montée va se faire [...] par un large  
 DEF.ART climb go.PRS.3SG REFL make.INF via INDF.ART wide  
 sentier [...] j’ai trouvé qu’on était monté plus que  
 path I.have.PRS.1SG find.PTCP that.3SG be.IPFV.3SG climb.PTCP more than  
 ce qu’on était descendu.  
 it that.3SG be.IPFV.3SG descend.PTCP

‘The climb will be made via a wide path; I found that we had climbed more than we had descended.’

In the context of the verb *monter*, the construction may incorporate a locative adverbial, e.g., the pronoun *y* in (24), or it may lack one (25).

**Quantitative evolution construction** The quantitative evolution construction corresponds to a derived meaning. In the case of *monter*, this construction can be observed with both ‘have’ and ‘be’ (contrarily to *passer*, where it only selects ‘be’), but it prefers ‘have’.

Consider (26) and (27):

- (26) [...]le rendement de l’emprunt d’État à 10 ans a  
 DEF.ART yield of. DEF.ART.bond of.government at 10 years have.PRS.3SG  
 monté de 3,6% à 3,9%.  
 rise.PTCP from 3,6% to 3,9%

‘The yield of the 10-year government bond rose from 3.6% to 3.9%.’

- (27) En 4 ans, les profits du Dow sont restés stables  
 in 4 years DEF.ART profits of. DEF.ART D. be-PRS.3SG remain.PTCP stable  
 mais le DOW est monté de 43%.  
 but DEF.ART D. have.PRS.3SG rise.PTCP by 43%.

‘In 4 years, Dow profits have remained stable, but the DOW has risen by 43%.’

In (26) and (27), the quantitative evolution construction is evident through the structure ‘from 3,6% to 3,9%,’ denoting a quantitative increase. However, even in sentences where quantity is not explicitly expressed, the quantity evolution meaning may occur.

Similar to *passer*, the construction type with *monter* also influences the meaning and the type of subject. When the verb is used in its spatial motion construction, the associated meaning is primary, and the preferred auxiliary is ‘be’. In this scenario, the canonical subject is human (28):

- (28) Quelques personnes sont montées, mais il restait plein  
 a.few people be.PRS.3PL go.up.PTCP.F.PL but it remain.IPFVS.3SG plenty  
 de places.  
 of seats

‘A few people went up, but there were still plenty of seats left.’

The verb used in the quantitative evolution construction conveys a derived meaning. In such instances, the subject is typically non-human, as exemplified by (29). However, human subjects can also appear with derived meaning, as illustrated by (30).

- (29) [...] le Dow Jones a monté de 0,45% et le Nasquad  
 DEF.ART D.J. have.PRS.3SG rise.PTCP by 0,45% and DEF.ART N.  
 de 0,32%  
 by 0,32%

‘[...] the Dow Jones rose by 0.45% and the Nasdaq by 0.32%.’

- (30) [...] regarder le nombre de visiteurs, avec à chaque connexion  
 watch.INF DEF.ART number of visitors with at every connection  
 le suspense de savoir si les visiteurs ont monté  
 DEF.ART DEF.ART of know.INF if DEF.ART visitors have.PRS.3PL rise.PTCP  
 ou s’ils ont chuté.  
 or if-they have-PRS.3PL fall-PTCP

‘[...] watch their visitor numbers, with the suspense of knowing at each connection whether visitors have risen or fallen.’

In (30), the subject is human but refers to the *number* of visitors rather than the visitors themselves. Therefore, it is human by metonymy, but the actual subject is ‘number’. The data

clearly indicate a preference for ‘be’ with primary meanings and ‘have’ with derived meanings due to the construction’s type.

It should be noted that spatial motion constructions, i.e., with their primary meaning, can also convey a figurative interpretation, as seen in (31).

- (31) Pendant la terreur c’est la période où Robespierre  
 during DEF.ART terror it.be.PRS.3SG DEF.ART period when R.  
 est monté au pouvoir.  
 be.PRS.3SG climb.PTCP at.DEF.ART power

‘During the terror is the period when Robespierre rose to power’.

In (31), a spatial movement is associated with a figurative meaning. The clear presence of a spatial motion construction ‘climb at location X’ categorizes it nevertheless as an instance of primary meaning, which is further emphasized by the involvement of a human subject.

Consider (32) and (33):

- (32) Le niveau des eaux a monté.  
 DEF.ART level of. DEF.ART waters have.PRS.3SG rise.PTCP

‘The water level has risen.’

- (33) [...] je trouve que la température est montée haute  
 I think.PRS.1SG that DEF.ART temperature be.PRS.3SG rise.PTCP.F.SG high  
 pas loin du rouge et j’ai perdu  
 NEG far from. DEF.ART red and I.have.PRS.1SG loose.PTCP  
 du liquide [...].  
 from. DEF.ART liquid

‘[...] I think the temperature has risen to near red and I’ve lost liquid [...].’

In (32) and (33), the subject represents a measurable entity, which points towards the derived meaning of quantitative evolution. The subject’s nature can therefore reinforce the derived interpretation.

Conversely, the more human a subject is, the more it tends to align with a primary meaning. Examples (34) and (35) illustrate this distinction:

(34) Je suis monté à 2000 à l'époque... et je savais pas  
 I be-PRS.1SG climb.PTCP to 2000 at.DEF.ART.time and I know.IPFV.1SG NEG

vraiment jouer [...] on aurait monté plus haut  
 really play.INF 3SG have-COND.3SG go.up.PTCP more high

'I climbed up to 2000 at the time... and I didn't really know how to play [...] we could have gone up higher.'

(35) Mon niveau d'endorphine est monté très haut.  
 my level of.endorphin be.PRS.3SG go.up.PTCP very high

'My endorphin level went very high.'

(34) exemplifies the spatial motion construction, where *monter* conserves its primary meaning, despite the figurative interpretation (cf. (31) 'rising to power'). This determination stems from the very nature of the construction, including the presence of the goal, and is strengthened by human nature of the subject. Again, the spatial motion construction is not incompatible with a figurative meaning.

Conversely, when the same construction occurs with a non-human subject, corresponding moreover to a measurable entity, as in (35), it unequivocally represents a derived meaning. The very nature of the construction implies something quantifiable, such as the level of endorphins, and conveys the derived meaning of quantitative evolution.

The primary factor in selecting the auxiliary is the type of construction, spatial motion or quantitative evolution. This distinction is correlated to the distinction between derived and primary meanings, since the spatial motion construction coincides with the primary meaning (both for *monter* and *passer*). Other factors, such as the nature of the subject (human or non-human, agentive or non-agentive), may also influence the choice of auxiliary, potentially specifying the type of construction and/or meaning conveyed.

***Human and non-human subjects*** In accordance with the above analysis, the choice of auxiliary depends indirectly on the [ $\pm$ HUMAN] nature of the subject because of the construction they are embedded in: human subjects demonstrate a preference for 'be', whereas non-human subjects tend to favor 'have.'

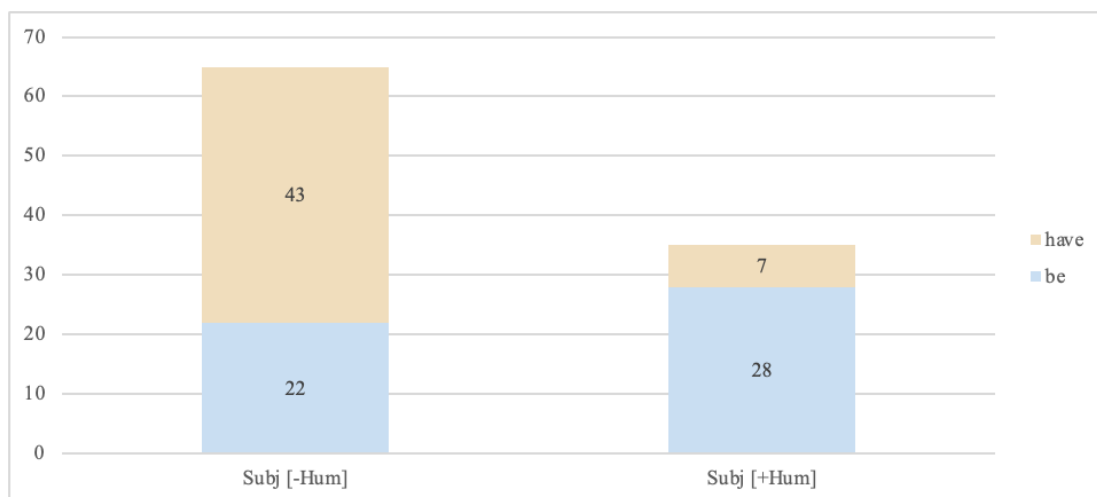


FIGURE 5.5 - Correlation between the nature of the subject and the selection of ‘have’ and ‘be’.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	22	43	65
Subj [+Hum]	28	7	35
<b>Total</b>	50	50	100

TABLE 5.6 – Number of ‘have’ and ‘be’ occurrences according to the factor [± HUMAN].

As observed in the case of *passer*, the type of construction emerges as a crucial factor. The correlation between the spatial motion construction and ‘be’, typically involving human subjects, illustrates this point. Conversely, the quantitative evolution construction, conveying a derived meaning, exhibit flexibility in selecting both ‘have’ and ‘be’, albeit with a preference for ‘have’.

Building upon the analysis of *passer* and the previous section, it can be inferred that when the subject is human, the verb tends to be utilized in a spatial motion construction. Consider (36):

- (36) La dernière fois que j’y suis monté, il y  
DEF.ART last time that I.LOC be.PRS.3SG climb.PTCP 3SG LOC  
 avait des petits stands de vendeur.  
have.IPFVS.3SG of. DEF.ART small stands of vendor

‘La dernière fois que j’y suis monté, il y avait des petits stands de vendeur.’

However, there are still some occurrences with ‘have’. In fact, the spatial motion construction favors a human subject, but human subjects may occur in quantitative evolution constructions.

With a few exceptions, the 7 tokens with ‘have’ mainly exhibit a derived meaning or a metonymical subject, as seen in example (37).

- (37) Le ton a monté et j’ai alors  
 DEF.ART tone have.PRS.3SG rise-.PTCP and I.have.PRS.1SG then  
 constaté que celui-ci maîtrisait mal le français.  
 realize.PTCP that this.one control.IPFVS.3SG badly DEF.ART french

‘The tone rose and I realized that he had a poor command of French.’

In (37), the meaning is derived (as it does not represent the primary spatial motion meaning), and the subject is metonymical as it stands for the person.

When considering non-human subjects, there is a higher frequency of ‘have’, indicating a notable preference for this auxiliary. However, it is essential to note that ‘be’ still occurs, although less frequently, with only 22 instances, compared to 43 of ‘have’.

Consider (38) and (39):

- (38) Les prix ont monté, ce n’est pas la même  
 DEF.ART prices have.PRS.3PL rise.PTCP it not.be.PRS.3SG NEG DEF.ART same  
 chose.  
 thing

‘Prices have increased, it’s not the same thing.’

- (39=27) En 4 ans, les profits du Dow sont restés  
 in 4 years DEF.ART profits of. DEF.ART D. be.PRS.3SG remain.PTCP  
 stables mais le DOW est monté de 43%.  
 stable but DEF.ART D. have.PRS.3SG rise.PTCP by 43%.

‘In 4 years, Dow profits have remained stable, but the DOW has risen by 43%.’

In this context, the auxiliary ‘be’ remains grammatically possible, as it aligns with the intended meaning of ‘rise’. However, ‘have’ emerges as the preferential choice. This preference is correlated to the derived meaning of the verb.

**Agentivity or lack of agentivity** The factor of agentivity indirectly correlates with the auxiliary selection, since it is conditioned the construction type. Specifically, when the subject is agentive or an internal cause, the preferred auxiliary is ‘be’. This preference arises because

an agentive or internal cause subject is more likely to be associated with a spatial motion reading which is the primary meaning of *monter*.

Conversely, non-agentive and non-internal cause subjects tend to favor ‘have’, indicating instances of quantitative evolution or derived meanings constructions, as Figure 5.6 illustrates.

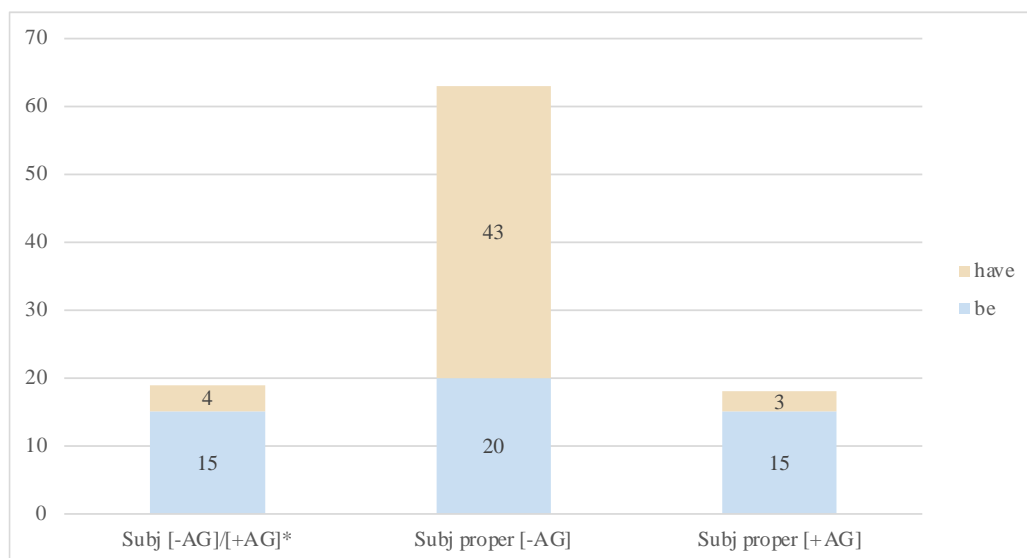


FIGURE 5.6 - Correlation of agentivity, internal cause, and lack of agentivity and ‘have’ and ‘be’.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-AG]/[+AG] *	15	4	19
Subj proper [-AG]	20	43	63
Subj proper [+AG]	15	3	18
<b>Total</b>	50	50	100

TABLE 5.7 – Total occurrences relative to ‘have’ and ‘be’ according to the factor [ $\pm$  Internal Cause] and [ $\pm$  Agentive].

Agentive subjects are exclusively human in the sample data and demonstrate a preference for ‘be’. They connect to the verb in its primary meaning of spatial motion, with one exception. The substantial predominance of ‘be’ aligns with the behavior of the corresponding verb in Italian, which also typically selects ‘be’ in its intransitive form.

Now, let us delve into a unique and particularly intriguing instance where both ‘have’ and ‘be’ auxiliaries are employed within the same sentence (39).

(39) Comme je suis curieux, je suis monté à l'arbre [...].  
 since I be.PRS.1SG curious I be.PRS.1SG climb.PTCP at.DEF.ART. tree

Et j'ai monté, j'ai monté...  
 and I.have.PRS.1SG climb.PTCP I.have.PRS.1SG climb.PTCP

'Since I'm curious, I climbed the tree. [...] And I climbed, I climbed...'

This sentence offers several noteworthy observations. In the initial segment where the auxiliary 'be' is employed, the emphasis is placed on the resultant state (*à l'arbre*). However, in the latter part of the sentence, the repetition suggests an ongoing process, hence the use of the auxiliary 'have'. Consequently, it can be inferred that while *monter* in its spatial construction typically favors 'be', the use of 'have' may be warranted when denoting an ongoing process without a specified endpoint.

This variation underscores two key points: firstly, although 'be' is the predominant auxiliary in such contexts, the 'be'/'have' selection is also related to aspect and aligns with a distinction between a 'resultativity state' vs an 'ongoing process' interpretation. 'Be' is preferred for a resultative state, whereas 'have' may be employed to highlight the process. Moreover, while 'be' remains the more prevalent auxiliary, it is noteworthy that 'have' exhibits greater frequency among intransitive verbs in French compared to Italian.

With *monter*, occurrences featuring internal cause subjects also tend to favor 'be' (40), although 'have' occurs sporadically (41).

(40) Le ton est monté et s'est terminé par un  
 DEF.ART tone be.PRS.3SG rise.PTCP and REFL.be-PRS.3SG end.PTCP with IND.F.ART

échange de coups devant l'établissement scolaire.  
 exchange of blows in.front.of DEF.ART.building educational

'The tone rose and ended with an exchange of blows in front of the school.'

(41=37) Le ton a monté et j'ai alors  
 DEF.ART tone have.PRS.3SG rise.PTCP and I.have.PRS.3SG then  
 constaté que celui-ci maîtrisait mal le français.  
 realize.PTCP that this.one control.IPFV.3SG badly DEF.ART french

'The tone rose and I realized that he had a poor command of French.'

(40) and (41) feature the same subject: 'tone'. This subject is considered human due to metonymy, representing the person who employs the tone. However, it cannot be regarded as voluntary, as the action pertains to the 'tone' itself rather than the person using it. As already

demonstrated in the section dedicated to Italian results, the internal cause factor can lead to significant variance in the selection between ‘have’ and ‘be’.

The variance resulting from the parameter of ‘internal cause’ can stem from a combination of different elements. On one hand, some internal cause subjects with *monter* are human and engaged in a spatial motion construction that primarily selects ‘be’, irrespective of the subject’s human or non-human nature. In a canonical spatial motion construction, the subject is typically agentive (as the movement is voluntary), but the context can bring a nuance, as in (42):

(42) Le           pilote   aura            monté       aussi   haut   qu’il    a  
DEF.ART   pilot   have.FUT.3SG   climb.PTCP   so   high   that.he   have.PRS.3SG

pu [...].  
can.PTCP

‘The pilot climbed as high as he could [...].’

In (42), the subject could potentially have been agentive. However, the subject is categorized as an internal cause because the action is not entirely voluntary; external limitations may have influenced the pilot’s ascent.

At the same time, some non-human subjects may select ‘have’ when they are internal causes due to their reference to a human feature in (43). Human subjects are typically associated with spatial motion constructions with primary meanings (and, therefore, cooccur with ‘be’).

However, in a construction with a derived meaning with a subject (indirectly) referring to humans, the strong association of agentivity and internal causality with ‘have’ may emerge (agentivity and internal causality are more readily associated with human features).

Consider a comparison between (43) and (44):

(43) Mais   la            polémique   a                    monté,    jusqu’à ce que [...]   le  
But   DEF.ART   discussion   have.PRS.3SG   rise.PTCP   until.that   DEF.ART

Conseil   représentatif   des            institutions   juives   de   France  
C.           R.                   of. DEF.ART   institutions   juives   of   France

condamne            fermement    l’invitation [...]  
condemn.PRS.3SG   strongly   DEF.ART.invitation

‘But the discussion escalated, until the Conseil représentatif des institutions juives de France strongly condemned the invitation.’

(44) [...] une seconde après (et c'était trop tard) la colère  
 INDF.ART second after and it.be.PRS.3SG too late DEF.ART anger

est montée en moi.  
 be.PRS.3SG rise.PTCP.F.SG in me

‘[...] a second later (and it was too late), anger rose in me.’

Both (43) and (44) depict a non-human subject considered as an internal cause. In (43), the controversy arises from a human referent, as if this referent has provoked it, whereas in (44), the rising anger is not controlled. The metaphorization of emotions as undergoing movement within a container can contribute to the use of ‘be’, characteristic of spatial motion constructions.

To conclude, non-agentive subjects typically cooccur with ‘have’. This tendency can be attributed to the fact that they are associated with the quantitative evolution construction rather than with the spatial motion construction of the verb *monter*. Consider (45) and (46):

(45=30) [...] regarder le nombre de visiteurs, avec à chaque connexion  
 watch.INF DEF.ART number of visitors with at every connection

le suspense de savoir si les visiteurs ont monté  
 DEF.ART DEF.ART of know.INF if DEF.ART visitors have-PRS.3PL rise.PTCP

ou s'ils ont chuté.  
 or if,they have.PRS.3PL fall.PTCP

‘[...] watch their visitor numbers, with the suspense of knowing at each connection whether visitors have risen or fallen.’

(46) Mais au moment où elles sont mises sur le  
 but at.DEF.ART moment when they be.PRS.3PL put.PTCP.F.PL on DEF.ART

marché, la valeur du bitcoin a monté [...]  
 market DEF.ART value of.DEF.ART bitcoin have.PRS.3SG rise.PTCP

‘But by the time they are put on the market, the value of bitcoin has risen [...].’

(45) features a human non-agentive subject, where the noun ‘visitors’ denotes the quantity, indicating the increase or decrease in numbers on the site. The example presents a quantitative construction similar to (46), which epitomizes the typical instance of derived meaning. Once

more, the agentivity factor aligns with other constructions: non-agentive subjects tend to opt for ‘have’ in sentences with derived meanings. Although ‘have’ predominates, instances of ‘be’ are still possible. Witness (47) and (48):

(47) Dans ma chambre, il fait en permanence dans les 22-23°C  
 in my bedroom it do.PRS.3SG in permanence in DEF.ART 22-23°C

voire plus (cet été je suis monté à 30°) [...].  
 even more this summer I be.PRS.1SG go\_up-PTCP to 30°

‘In my bedroom, it’s always 22-23°C or even higher (this summer I went up to 30°C) [...].’

(48) [...] cela ne gênait nullement la banque d’appliquer cet  
 this not bother.IPFV.3SG in.no.way DEF.ART bank to.apply.INF this

index qui était monté à >10% [...].  
 index that be.IPFV.3SG rise.PTCP to >10%.

‘[...] the bank had no problem applying this index, which rose to >10%[...].’

(48) displays a quantitative evolution construction with a human subject referring to the temperature measure (‘this summer I went up to 30°’ should be ‘this summer the temperature in my room went up to 30°’). However, the presence of a complement introduced by *à*, be it non locative, indicating the endpoint of the evolution, might have prompted the usage of ‘be’. The same configuration is offered in (47), where the quantitative evolution construction can be inferred from the nature of the subject, non-human and corresponding to a quantifiable entity and the auxiliary is nevertheless ‘be’.

In addition to the construction and the conveyed meaning, there are other secondary factors. Human subjects tend to favor ‘be’, as they are more commonly associated with the spatial motion construction, while non-human subjects tend to prefer ‘have’, since they are more frequently found in quantitative-derived meaning constructions. The semantic attributes of the subjects are linked to the types of constructions in which they are found. Similarly, syntactic contexts, such as adverbials, correlate with ‘have’ or ‘be’, on the basis of the constructions in which they are employed.

**Adverbials** Given the usage of the verb *monter* within quantitative evolution and spatial motion constructions, it is often accompanied by specific adverbials. Notably, these adverbials frequently include those denoting quantity, space, and time adverbials.

**QUANTITY ADVERBIALS** In the context of quantity adverbials, both ‘have’ and ‘be’ are observed within the quantitative evolution construction associated with the verb *monter*. However, there is a notable preference for ‘have’ in such instances with non-human subjects. It is important to note that this preference is not inherently linked to the quantity adverbial itself, but rather to the quantitative evolution construction in which it is embedded.

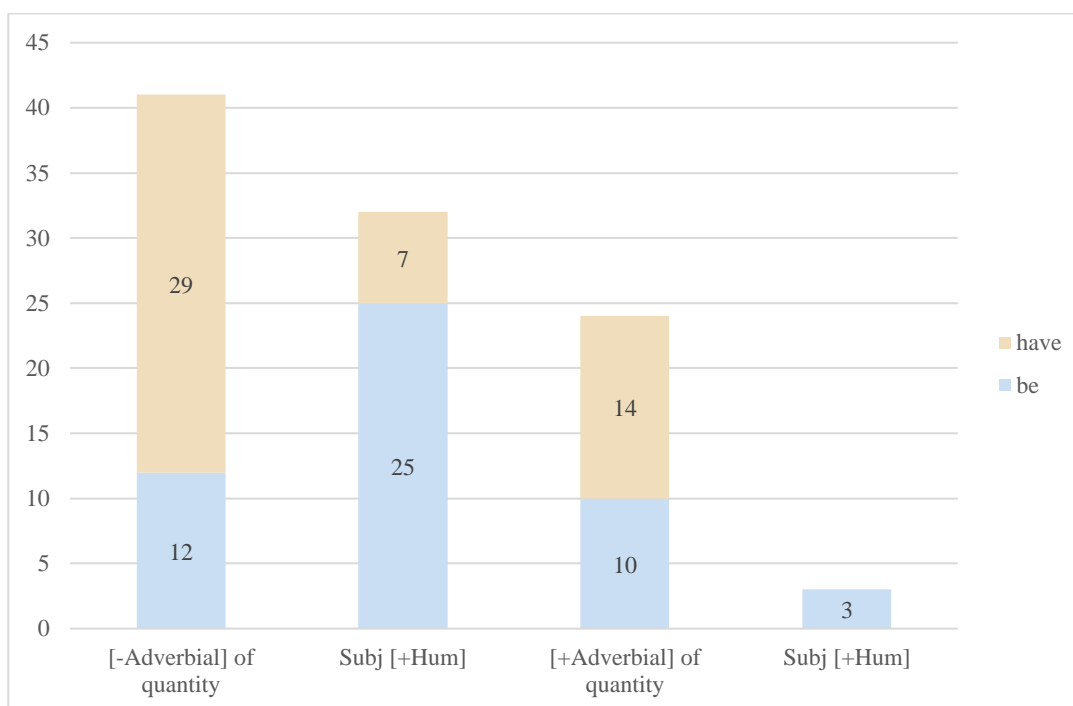


FIGURE 5.7 - Correlation between ‘have’ and ‘be’ with quantitative adverbials and human subjects.

Factors	Data ‘be’	Data ‘have’	Total
<b>[-Adverbial] of quantity Subj [-Hum]</b>	12	29	41
<b>[-Adverbial] of quantity Subj [+Hum]</b>	25	7	32
<b>[+Adverbial] of quantity Subj [-Hum]</b>	10	14	24
<b>[+Adverbial] of quantity Subj [+Hum]</b>	3	/	8
<b>Total</b>	50	50	100

TABLE 5.8 - Total occurrences relative to ‘have’ and ‘be’ according to [± Adverbial of quantity and ± HUMAN]

(49) illustrates the selection of ‘have’ for the quantitative evolution construction:

- (49) La bourse a monté de près de 15% aux  
DEF.ART stock market have.PRS.3SG rise.PTCP by close.to 15% at.DEF.ART
- troisième et quatrième trimestres.  
third and fourth quarters

‘The stock market rose by nearly 15% in the third and fourth quarters.’

The auxiliary ‘be’ is not precluded, yet ‘have’ is notably favored. The adverbial of quantity predominantly accompanies quantitative evolution constructions. The majority of sentences featuring this adverbial exhibit a non-human subject, with however three exceptions.

An example is given in (50):

(50) Sur mon mini capital de 300€ j’étais monté en quelques  
 On my mini capital of 300€ I.be.IPFV.1SG go\_up.PTCP in a.few  
 jours.  
 days

‘On my mini-capital of 300€ I had risen to 500€ in a few days.’

In (50), the subject has risen in price, but there is once again an edge and an increase transferred to a human subject.

**LOCATIVE ADVERBIALS** The sample data suggest a predominance of ‘be’ when paired with locative adverbials denoting a denoting a place that is the target or end point of the movement (e.g., *on y était monté* – ‘we went up there’), especially in conjunction with human subjects. This consistent preference for ‘be’ aligns with instances featuring human subjects, which typically correspond to the spatial motion construction and convey the primary meaning. The occurrences featuring both ‘be’ and ‘have’ are illustrated in Figure 5.8.

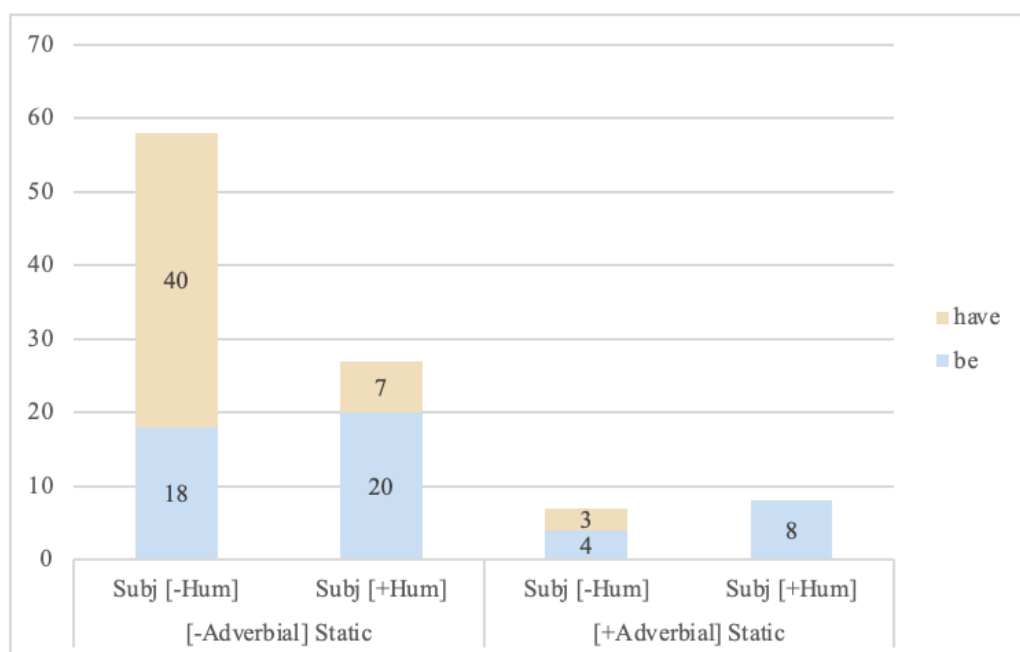


FIGURE 5.8 - Correlation between ‘have’ and ‘be’ with locative adverbials and human subjects.

Factors	Data 'Be'	Data 'Have'	Total
[-Adverbial] static Subj [-Hum]	18	40	58
[-Adverbial] static Subj [+Hum]	20	7	27
[+Adverbial] static Subj [-Hum]	4	3	7
[+Adverbial] static Subj [+Hum]	<b>8</b>	/	8
<b>Total</b>	50	50	100

TABLE 4.9 – Total occurrences relative to ‘have’ and ‘be’ according to [ $\pm$  Locative adverbial] and [ $\pm$  HUMAN].

As indicated in Table 4.9, there is a 100% occurrence of ‘be’ with human subjects in conjunction with locative adverbials. This consistent preference for ‘be’ with human subjects can be attributed to the association with the spatial motion construction and the primary or literal meaning.

In contrast, data involving non-human subjects exhibit more variability, as these subjects are not typically used with the primary or literal meaning of the verb, thus allowing for a larger usage of ‘have’. However, instances of ‘be’ with non-human subjects may still arise, possibly influenced by the spatial motion construction.

Consider example (51):

(51)	[...] pendant	les	gouvernements	précédents	l'économie	se
	during	DEF.ART	government	previous	DEF.ART.economy	REFL
	portait bien [...]	en	aggravant	la	dette, on	voit
	do.well.IPFV.3SG	in	increasing	DEF.ART	debt 3SG	see.PRS.3SG
	où	elle	est	montée !		
	where	3SG	be.PRS.3SG	go_up.PTCP.F.SG		

‘[...] and during previous governments the economy was doing well [...] by increasing the debt, we can see where it went!’

The conveyed meaning is a quantitative evolution, but the underlying construction remains strictly linked to locative. This inherent spatial motion nature may play a significant role in the selection of the auxiliary ‘be’.

**TEMPORAL ADVERBIALS** The temporal adverbial (e.g., *le lendemain* ‘the following day’) does not appear to exert influence over the selection of the auxiliary verb. Despite its frequent occurrence, establishing a temporal context, the distribution of auxiliaries does not vary in function of this parameter. In its presence, there are 8 sentences featuring ‘have’ and 9 featuring ‘be’.

In summary, the verb *monter* offers two distinct constructions, conveying distinct meanings: spatial motion and quantitative evolution. Within the spatial motion construction, where *monter* has its primary meaning, the preferred auxiliary is ‘be’. Conversely, in the quantitative evolution construction, which conveys a derived meaning, ‘have’ emerges as the favored auxiliary, although ‘be’ may still be used in specific contexts.

Human subjects, acting as agents, exhibit a propensity for selecting ‘be’ when engaging with *monter* in its spatial motion meaning of ‘climb’. Metaphorical interpretations are more commonly associated with non-human subjects, thereby contributing to a higher incidence of ‘have’. The affinity between [+HUMAN] / [+AGENTIVE] and ‘be’ on one hand, and between [-HUMAN] / [-AGENTIVE] and ‘have’ on the other, results from the fact that the construction is paramount in the selection of the auxiliary, with the subject features being merely consequential.

Furthermore, a distinction emerges between processive meaning and resulting state meaning: processes tend to favor ‘have’, whereas resulting states lean towards ‘be’. Table 6.10 provides a comprehensive overview of these primary factors.

Construction	Human trait	Agentivity	Auxiliary
Spatial motion	Human	Agentive   Internal cause →	Be
Quantitative	Non-human	Not agentive →	Have

TABLE 6.10 – Primary factors influencing the choice of ‘have’ and ‘be’ for *monter*.

Table 6.10 provides insight into the primary factors influencing the selection of the auxiliary verb. Within spatial constructions, the subject typically assumes a human and agentive role, closely linked to the primary meaning, thus favoring ‘be’ as the preferred auxiliary.

In contrast, quantitative evolution constructions often feature a non-human, non-agentive subject, associated with derived meanings, thereby exhibiting a preference for the auxiliary ‘have’.

#### 4.1.3 *Baisser*

*Baisser*, translated as ‘go down’ when intransitive, is a verb that frequently appears with ‘have’. In the FrTenTen corpus (17), 95% of instances use ‘have’ and 5% use ‘be’ when the filter ‘no noun in the first five tokens right’ is applied, with 6,356 occurrences of ‘have’ and 326

occurrences of ‘be’. *Baisser* is a labile verb, but this study will focus solely on its intransitive construction.

**Quantitative evolution construction and meaning** In comparison to *monter*, which encompasses both a literal meaning of spatial motion (often favoring the auxiliary ‘be’) and a derived meaning indicating a quantitative change (frequently associated with ‘have’), *baisser*, despite also having a literal meaning of spatial motion, is predominantly employed in its derived or quantitative sense within this corpus, as illustrated in (52) and (53).

(52) Le niveau de la mer a baissé [...]  
 DEF.ART level of. DEF.ART sea have.PRS.3SG decrease.PTCP  
 ‘The sea level has fallen [...]

(53) Oui, les prix ont baissé depuis peu [...]  
 yes DEF.ART prices have.PRS.3PL decrease.PTCP since a.bit  
 ‘Yes, prices have recently fallen [...]

In both constructions, the expression is quantitative as both subjects are measurable in quantity. However, (52) presents a more literal interpretation. The subjects co-occurring with this verb are non-human.

**Non-human subjects** In the sample, 50 instances of ‘have’ and 50 instances of ‘be’ have been meticulously chosen to delve into the factors influencing the selection of the auxiliary verb. Notably, the data exclusively involve non-human subjects. Figure 4.9 illustrates an equal distribution of sentences featuring ‘have’ and ‘be.’

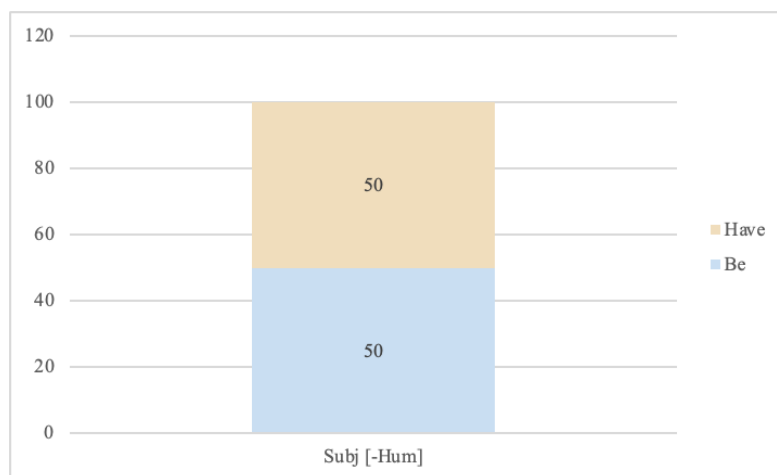


FIGURE 4.9 - Correlation between the use of ‘have’ and ‘be’ and the human trait.

Factors	Data 'Be'	Data 'Have'	Total
Subj [-Hum]	50	50	100
<b>Total</b>	50	50	100

TABLE 4.11 – Total occurrences relative to ‘have’ and ‘be’ according to [ $\pm$  HUMAN]

As evidenced by both the Figure and the Table, all sentences involve non-human subjects and employ ‘have’ and ‘be’. Considering that the only the quantitative evolution construction is attested, it can be inferred that both ‘have’ (54) and ‘be’ (55) are possible options for this construction.

- (54) [...] les prix de vente ont baissé de 8% [...].  
DEF.ART price of sale have.PRS.3PL decrease.PTCP by 8%

‘[...] sales prices decreased by 8% [...].’

- (55) Heureusement que le prix est baissé [...]  
Fortunately that DEF.ART price be.PRS.3SG decrease.PTCP

‘Fortunately, the price has decreased [...].’

Despite both ‘have’ and ‘be’ being possible, the larger FrTenTen (17) corpus shows a notable preference for ‘have’.

Another aspect to consider is that instances of ‘be’ can be analyzed in two ways: either as an auxiliated form of the intransitive construction or as a passive voice of the transitive construction. A case in point is (56): the former analysis can be glossed as ‘the costs will drop’, whereas the analysis as a passive voice corresponds to ‘the costs will be dropped’.

- (56) Bref les coûts seront baissés.  
in.short DEF.ART costs be.FUT.3PL drop.PTCP.MPL

‘In short, the costs will drop/will be dropped.’

Both analyses are equally viable, and deciding between them is impossible without considering the context: this ambiguity arises from the nature of *baisser*, which is a ‘labile’ verb, a verb characterized by causative alternation, with a transitive construction, e.g. *Les banques ont baissé les prix de vente*, and an intransitive construction, containing as subject the direct object of the corresponding transitive construction : *Les prix de vente ont baissé*.

As demonstrated by Buchard & Carlier (2008: 2429), the intransitive construction of these verbs may be compatible with both ‘have’ and ‘be’ as auxiliary verbs in the *passé composé*. This

opposition reflects an aspectual distinction between a processive reading and a resultative reading, close to adjectival predication.

**Agentivity and lack of agentivity** The potential ambiguity between the intransitive and passive constructions increases by the massive predominance of non-agentive subjects, which are also the prototypical subjects of passive constructions. Figure 4.10 represents the frequency of non-agentive subjects with *baisser*.

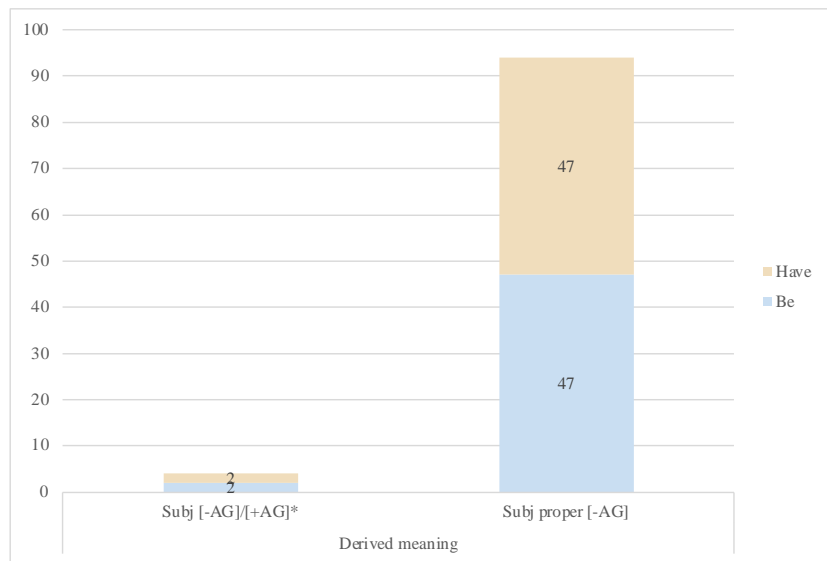


FIGURE 4.10 - Correlation between the selection of ‘have’ and ‘be’ and the factor of agentivity in derived meanings

Factors	Data ‘Be’	Data ‘Have’	Total
Derived meaning Subj [-AG]/[+AG] *	2	2	4
Derived meaning Subj proper [-AG]	47	47	94
<b>Total</b>	<b>49</b>	<b>49</b>	<b>98</b>

TABLE 4.12 – Total occurrences relative to ‘have’ and ‘be’ according to [± Internal Cause] in derived meanings.

Table 4.12 demonstrates that the majority of occurrences (94 tokens) involve non-agentive subjects. Witness the following examples:

- (57) Ce taux est baissé à 10% pour les travaux[...]  
 This rate be.PRS.3SG drop.PTCP to 10% for DEF.ART work

d’amélioration  
 of.improvement

This rate has dropped to 10% for improvement work.

(58) Le taux de mortalité [...] avait baissé de 50% [...]  
 DEF.ART rate of mortality have.IPFV.3SG drop.PTCP by 50%

‘The mortality rate [...] had fallen by 50%.’

Both sentences exhibit the subject ‘rate’, which readily aligns with the derived meaning of *baisser*.

Instances with internal cause subjects are sparse. This is due to the fact that *baisser*, in its derived meaning ‘decrease’, evokes a process that does not necessarily involve control exerted by the subject. The subsequent examples feature subjects that imply human attributes:

(59) Je me rends compte que je ne tiens pas le rythme, ma  
 I REFL realize.PRS.1SG that I not keep.PRS.1SG NEG DEF.ART rhythm my

vitesse descente a baissé de 500 à 400 voir 300m/h.  
 speed descent have-PRS.3SG drop.PTCP from 500 to 400 or 300m/h.

‘I realize that I can’t keep up, my descent speed has dropped from 500 to 400 or even 300 meters per hour.’

(60) Quand la garde de sa proie est baissée, elle la  
 When DEF.ART guard of her prey be.PRS.3SG drop.PTCP.F.SG she it

prend par surprise.  
 take.PRS.3SG by surprise.

‘When her prey’s guard has dropped, she takes it by surprise.’

In (59), the subject is inanimate (*vitesse*), yet the allusion to a human participant can be inferred from the preceding segment of the sentence, *Je ne tiens pas le rythme*. In (60), a subject is observed exhibiting causative behavior: the term ‘guard’ implies control, but the verb *baisser* suggesting a loss of control.

**Adverbials** As for the adverbials co-occurring with *baisser*, the most frequent ones are the quantity and temporal adverbials.

**ADVERBIAL OF QUANTITY** Given that *baisser* is primarily employed in its derived meaning and in the quantitative evolution construction, adverbials of quantity such as ‘with 20%’ in *le stock de ventes a baissé de 20 %* are very frequent, occurring in 60 instances, as illustrated in Figure 4.11.

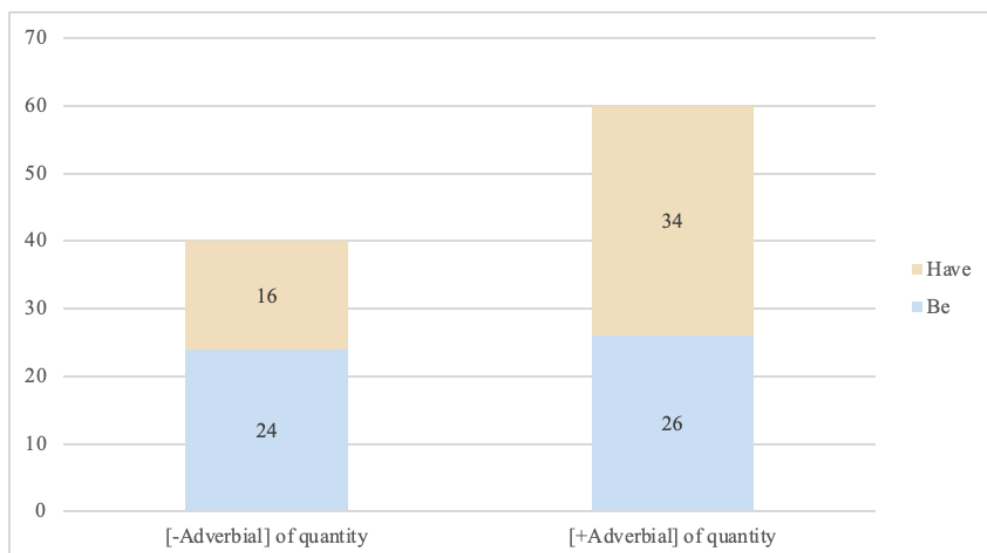


FIGURE 4.11 - Correlation between the choice of ‘have’ and ‘be’ and the adverbial of quantity.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-Adverbial] of quantity</b>	24	16	40
<b>[+Adverbial] of quantity</b>	26	34	60
<b>Total</b>	50	50	100

TABLE 4.13 – Total occurrences relative to ‘have’ and ‘be’ according to [± Adverbial of quantity]

The analysis reveals that both ‘have’ and ‘be’ are possible choices with *baisser* with an adverbial expressing quantity. Nonetheless, the prevalence of ‘have’ is apparent, and is confirmed by the broader FrTenTen (17) corpus. An example is provided in (61).

- (61) Les ventes de Toyota ont baissé de 8,7% sur un an  
DEF.ART sales of T. have.PRS.3PL decrease.PTCP by 8,7% year.on.year
- en février aux Etats-Unis.  
in february at.DEF.ART U.S.

‘Toyota sales decreased by 8.7% year-on-year in February in the U.S.’

**TEMPORAL ADVERBIALS** Occurrences with temporal adverbials (e.g., *in 2013*) show a more substantial preference for ‘have’ over ‘be’ (11 occurrences of ‘have’ compared to 5 occurrences of ‘be’), as depicted in Figure 4.12.

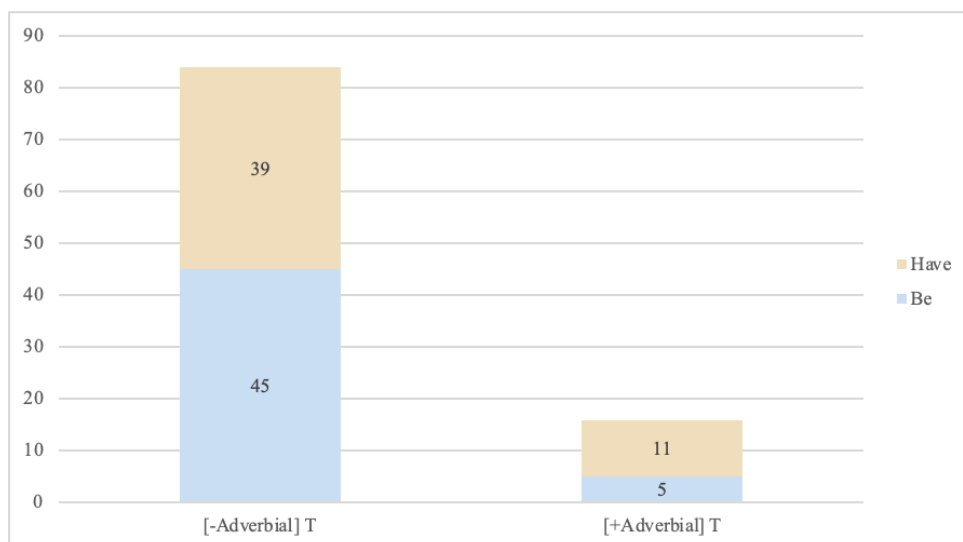


FIGURE 4.12 - Correlation between the choice of ‘have’ and ‘be’ and the temporal adverbial.

Factors	‘be’	‘have’	Total
<b>[-Adverbial] T</b>	45	39	84
<b>[+Adverbial] T</b>	5	11	16
<b>Total</b>	50	50	100

TABLE 4.14 – Total occurrences relative to ‘have’ and ‘be’ according to [± Temporal Adverbial]

Without the temporal adverbial, ‘be’ is slightly more prevalent, accounting for approximately 53.57% of instances, while ‘have’ occurs in about 46.43% of cases. However, with the inclusion of a temporal adverb, the frequency of ‘have’ significantly increases to approximately 68.75%, while ‘be’ decreases to around 31.25%. This disparity indicates a notably higher incidence of the auxiliary ‘have’ when a temporal adverb is present compared to ‘be’. It is possible to relate this tendency to an aspectual difference: the *passé composé* with ‘be’ more easily evokes the resulting state than the past tense with ‘have’, and this resulting state value is less compatible with specific dates.

In summary, the verb *baisser*, unlike the previously examined verbs *monter* and *passer*, does not behave as a conventional spatial motion verb because it is exclusively employed with non-human subjects. In the analyzed 100-sentences corpus, only the derived meaning of *baisser* was identified, with few exceptions.

Although both ‘have’ and ‘be’ auxiliaries are possible, an examination of various factors, including data from the larger FrTenTen (17) corpus and the possible passive interpretation, indicates that ‘have’ is the preferred auxiliary for *baisser* in its intransitive construction.

Table 4.15 delineates the most recurrent factors influencing the auxiliary selection:

Construction	Human trait	Agentivity	Auxiliary
Quantitative evolution	Non-human	Non-agentive ➡	Have
Quantitative - passive	Non-human	Non-agentive ➡	Be

TABLE 4.15 – Principal factors concurring in the selection of the auxiliary.

Table 4.15 provides a comprehensive overview of the factors influencing the selection of auxiliary verbs in various contexts. In contrast to *passer* and *monter*, *baisser* is in the sample data invariably used in its derived, quantitative meaning. When the construction has a quantitative interpretation, often accompanied by an adverbial of quantity (though not systematically), the subject typically lacks agency and tends to be non-human. In such scenarios, both auxiliaries tend to be used interchangeably, but ‘have’ emerges as the preferred auxiliary, considering that many occurrences of ‘be’ in the quantitative interpretation may instantiate a passive construction.

#### 4.1.4 *Descendre*

The verb *descendre* ‘go down’, ‘drop’ yields 8% of instances with ‘have’ when queried using CQL in the FrTenTen (17) corpus, filtered to exclude nouns within five words to the right, and 88% of occurrences with the auxiliary ‘be’. The occurrences found are 225 with ‘have’ and 2,718 with ‘be’.

While *descendre* is a labile verb, this study focuses solely on its intransitive construction. Unlike its Italian counterpart, which predominantly selects ‘be’ in its intransitive form, French exhibits more auxiliary alternation.

*Descendre* can be employed in both spatial motion and quantitative evolution constructions, but also in a construction at the intersection of spatial motion and quantitative evolution. Figure 4.13 represents the three constructions from a quantitative viewpoint.

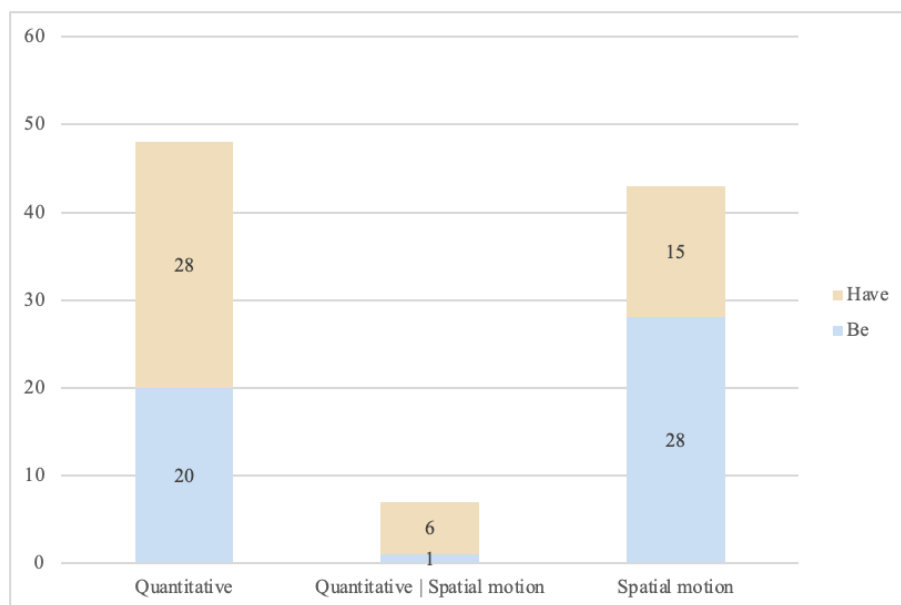


FIGURE 4.13 - Correlation between 'have' and 'be' and the construction's type.

Factors	Data 'Be'	Data 'Have'	Total
<b>Spatial motion</b>	28	15	43
<b>Quantitative</b>	20	28	48
<b>Quantitative   Spatial motion</b>	1	6	7
<b>Total</b>	49	49	98

TABLE 4.16 – Total occurrences relative to 'have' and 'be' according to the construction.

As depicted in table 4.16, two occurrences remain unidentified within the established categories, based on contextual analysis. The spatial motion construction shows a predilection for the verb 'be', whereas the quantitative evolution construction exhibits a preference for 'have'. Notably, the quantitative spatial motion form shows a distinct preference for 'have', constituting a subset of the broader quantitative category, as elaborated shortly.

This distinction is visually represented in the Figure to underscore the disparity from the canonical quantitative form.

***Spatial motion construction*** In its spatial motion construction, *descendre* permits the use of both auxiliary verbs, although 'be' is favored (62; 63).

- (62) Les habitants du quartier sont descendus de chez eux.  
 DEF.ART residents of. DEF.ART district be.PRS.3PL go.down.PTCP.M.PL from at.their.house

'The district's residents went down from their houses.'

(63) [...] Je suis descendu, il devait être 4h30.  
 I be.PRS.1SG go.down.PTCP 3SG must.IPFV.3SG be.INF 4h30.

‘[...] I went downstairs, it must have been 4h30.’

Spatial motion constructions with a primary meaning but interpreted figuratively are less frequent but not impossible, as exemplified in (66).

(64) [...] une jeune femme droguée qui est descendue très bas.  
 INDF.ART young woman on.drugs who be.PRS.3SG drop-PTCP.F.SG very low

‘[...] a young woman on drugs who has dropped very low.’

In (64), the interpretation is figurative, but the construction includes a locative adverbial, making it an instance of primary meaning.

**Quantitative evolution construction** When employed in the canonical quantitative evolution construction, *descender* is used in a derived meaning. Both ‘have’ and ‘be’ can be used, as exemplified in (65) and (66):

(65) [...] les ventes numériques d’albums ont descendu de 26%.  
 DEF.ART sales digital of.albums have.PRS.3PL drop.PTCP by 26%

‘[...] digital album sales have dropped by 26%.’

(66) Le chômage est descendu à 4,3% [...].  
 DEF.ART unemployment be.PRS.3SG drop.PTCP to 4,3%

‘Unemployment dropped to 4.3% [...]’.

Nevertheless, ‘have’ predominates in the quantitative evolution construction within the sample of 100 sentences. Additionally, *descendre* is a labile verb, and its ‘be’ construction can present ambiguity between an intransitive or a passive interpretation.

Despite the quantitative evolution construction favoring a non-human subject, as observed for other verbs analyzed above, *descendre* in this construction is also compatible with a human subject, as evidenced by examples (67) and (68).

(67) Il y a trois semaines j’ai descendu à 101.8 [Kg].  
 3SG LOC have-PRS.3SG three weeks I.have.PRS.1SG drop.PTCP to 101.8

‘Three weeks ago, I dropped to 101.8 [Kg].’

(68) L’an dernier, j’étais descendue au 91 [format].  
 DEF.ART -year last I.be.IPFV.1SG drop.PTCP.F.SG at. DEF.ART 91 [size].

‘Last year, I dropped to 91 [size].’

In (67) and (68), a decrease in weight is evident. Similarly, (68) describes a reduction in size that is comparable in meaning to the reduction in weight. All sentences featuring derived meanings with human subjects contextualize a decrease in weight or a situation similar to that in (68). The quantity can be expressed either through a quantity adverbial, as seen in (67; 68), or through an internal object, as illustrated in (69).

(69) [...] j’ai descendu aussi 60-62 kg.  
 I.have.PRS.1SG go.down.PTCP also 60-62 kg.

‘[...] I also went down to 60-62 kg.’

(69) is a clear instance of a pseudo-transitive structure which occurs both with the primary meaning of spatial motion and the derived meaning of quantitative evolution. It is therefore labeled as the ‘quantitative-locative construction’.

QUANTITATIVE | LOCATIVE Within the quantitative evolution construction with *descendre*, there exists a subset that can be defined as the quantitative-locative construction. In this subset, certain occurrences exhibit a quantitative evolution construction but differ by conveying a locative meaning. These instances often feature a pseudo-transitive construction with an internal object denoting the quantity. In such cases, ‘have’ is typically selected, as exemplified in (70).

(70) J’ai descendu 200m, puis remonté par le même  
 I.have.PRS.1SG go.down.PTCP 200m then climb.back.PTCP by DEF.ART same

itinéraire.  
 route

‘I descended 200m, then climbed back up by the same route.’

Analyzing FrTenTen (17), there is a significant disproportion between the uses of ‘have’ and ‘be’ with *descendre*. Some instances of ‘have’ can be attributed to this extended use of ‘have’ in French, especially where *descendre* is accompanied by an internal object expressing quantity, both with a derived or a primary meaning. The selection of ‘have’ when *descendre* has an

internal object can be accounted for by syntax, since the proximity of the internal object construction with the transitive construction.

Another important factor is the type of subject, which partly correlates with the type of construction: a spatial motion construction requires a human subject, whereas a quantitative evolution construction typically prefers (though not always, see (69)) a non-human subject.

**Human and non-human subjects** While both ‘have’ and ‘be’ can cooccur with both human and non-human subjects, ‘be’ is more frequent with human subjects and ‘have’ with non-human subjects, as Figure 4.14 illustrates.

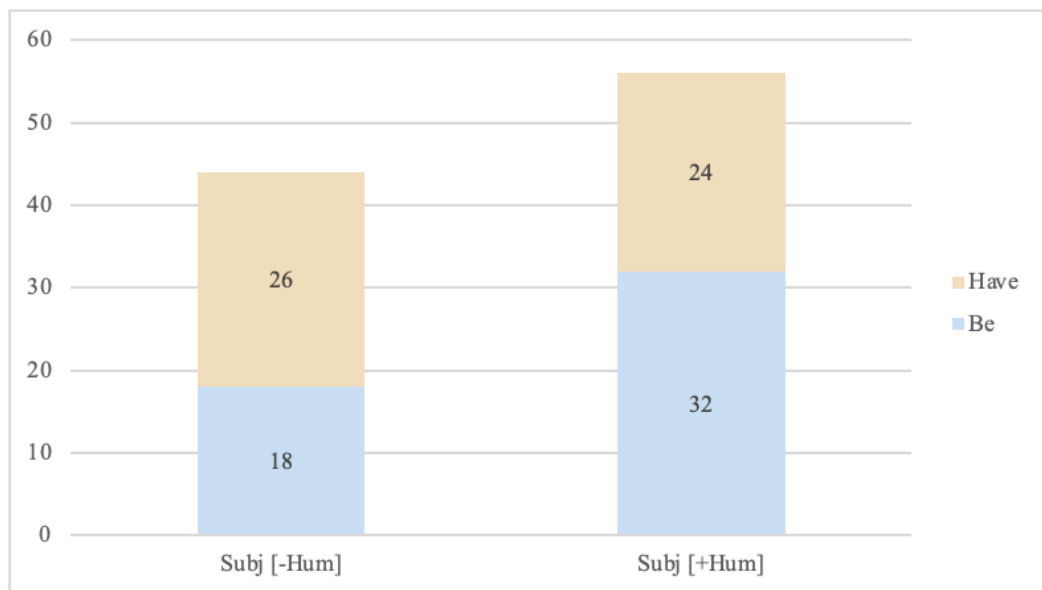


FIGURE 4.14 - Correlation between ‘have’ and ‘be’ and the human/non-human nature of the subject.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	18	26	44
Subj [+Hum]	32	24	56
Total	50	50	100

TABLE 4.17 – Total occurrences relative to ‘have’ and ‘be’ according to the [ $\pm$  HUMAN] factor.

The preference for ‘be’ with human subjects correlates with frequent association with the spatial motion construction, where *descendre* has its primary meaning. In contrast, the prevalence of ‘have’ with non-human subjects is grounded in its frequent association with the quantitative evolution construction, *i.e.* with a derived meaning. Consider examples (71) and (72):

(71) [...] Je suis descendu, il devait être 4h30.  
 I be.PRS.1SG go.down.PTCP 3SG must.IPFV.3SG be.INF 4h30.

‘[...] I went downstairs, it must have been 4h30.’

(72) [...] son taux de diabète a descendu [...].  
 his level of diabetes have.PRS.3SG go.down.PTCP

‘[...] his diabetes level went down [...].’

In (71), a human subject is associated with a spatial motion construction, representing a primary meaning. Conversely, in (72), a non-human subject is linked to a quantitative evolution construction, indicating a derived meaning.

Globally, non-human subjects tend to trigger the derived meaning interpretation, not only with ‘have’ (73) but also with ‘be’ (74).

(73) En 2011, 46,8% d’internautes passaient par un  
 in 2011, 46,8% of.internet.users go.IPFV.3PL through INDF.ART

ordinateur portable, ce pourcentage a descendu d’1%.  
 laptop this percentage have-PRS.3SG drop-PTCP by.1%

‘In 2011, 46.8% of Internet users used a laptop, a percentage that has now dropped by 1%.’

(74) Son taux d’oxygène dans le sang est descendu à 30%.  
 his level of.oxygen in DEF.ART blood be.PRS.3SG drop.PTCP to 30%

‘His blood oxygen level has dropped to 30%.’

The comparison between these examples is revealing because both instances represent a quantitative evolution construction and a derived meaning, yet (73) selects ‘have’ while (74) selects ‘be’.

The difference in the choice of the auxiliary can be attributed to the resultative interpretation that arises from the complement headed by the preposition *à* indicating the endpoint of the evolution in (74), whereas the expression of quantity ‘*de + quantity*’ can be subject to various interpretations.

Human subjects are predominantly associated with the spatial motion construction and the correlated primary meaning of *descendre*, and the construction strongly favors the use of ‘be’ (75):

(75) Je suis descendu très tôt pour aller chercher des  
 I be.PRS.1SG go\_down.PTCP very early to go.INF get.INF of. DEF.ART

journaux.  
 newspapers

‘I went down early to get some newspapers.’

In conclusion, the puzzle is quite intricate due to the simultaneous influence of various factors. Firstly, *descendre* is used in the spatial motion construction, which conveys a primary meaning and typically features a human subject. In this circumstance, the preferred auxiliary is ‘be’.

Secondly, *descendre* occurs in the quantitative evolution construction, which conveys a derived meaning and typically involves a non-human subject, showing a preference for ‘have’.

However, there are occurrences that fall between these two categories, where quantity constructions feature human subjects. These instances can convey both a primary meaning (e.g., *J’ai descendu 200m*) and a derived one (e.g., *J’ai descendu à 20kg*). With this construction, the preferred auxiliary is ‘have’, likely due to the proximity of the internal object construction (e.g., *J’ai descendu 200m*) to the transitive construction.

***Agentivity and lack of agentivity with human subjects*** Analyzing the features of agentivity or internal cause conveyed by the subject could provide insight into the complexity of the verb’s auxiliary selection. With human subjects, the subjects are especially internal cause and agentive. Occurrences with internal cause subjects show a clear preference for ‘be’, while agentive subjects show both ‘have’ and ‘be’. The presence of ‘have’ is not a contradiction as it refers to specific constructions, such as the pseudo-transitive, which will be detailed later.

Figure 4.15 delineates instances specifically featuring agentive and internal cause human subjects.

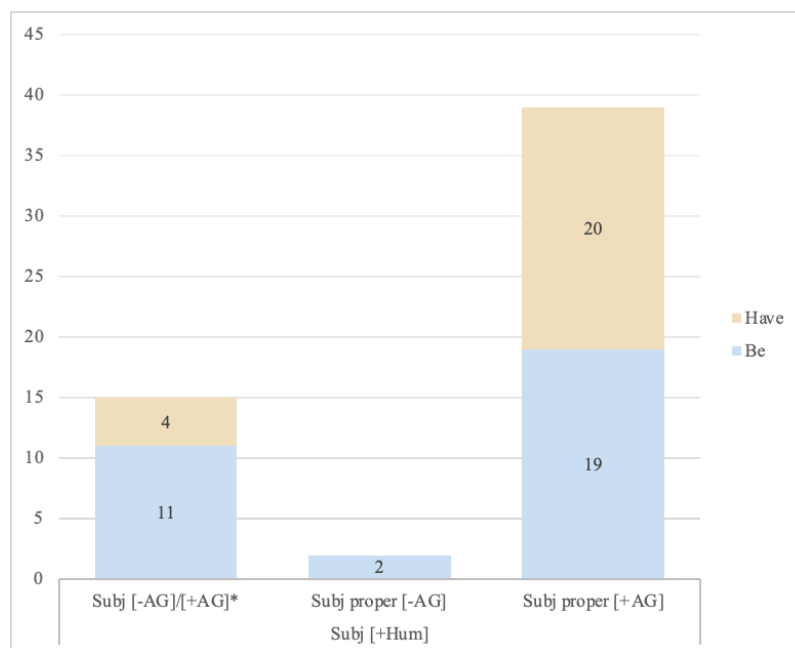


FIGURE 4.15 - Correlation between ‘have’ and ‘be’, internal cause and the human subject.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [+Hum] [-AG]/[+AG] *</b>	11	4	15
<b>Subj [+Hum] proper [-AG]</b>	2	/	2
<b>Subj [+Hum] proper [+AG]</b>	19	20	39
<b>Total</b>	32	24	56

TABLE 4.18 – Total occurrences relative to ‘have’ and ‘be’ according to [± HUMAN] and [± INTERNAL CAUSE].

As shown in Table 4.18, there are only two instances where a human subject is non-agentive. An example is provided in (76), where the abrupt drop in a medical result is attributed to a human and non-agentive subject. However, it is not the person who ‘went down’, but rather the level of what was analyzed during the test.

(76) J’ai fait mon test de jeûne, qui a été  
 I-have-PRS.1SG do.PTCP my test of fasting which have.PRS.3SG be.PTCP  
 arrêté [...] car je suis descendue à 0,30.  
 stop.PTCP because I be.PRS.1SG drop.PTCP.F.SG to 0,30

‘I did my fasting test, which was stopped [...] because I dropped to 0,30.’

A subject embodies the characteristic of an internal cause when it is accountable for an event without conscious intent. This is exemplified in sentences previously observed, such as weight loss (an individual may aim to shed weight but cannot predetermine the exact amount lost or guarantee success or he may also lose weight involuntarily).

Similarly, the subject can be construed as an internal cause in (77), where *descendre* retains its literal meaning of spatial motion:

- (77) Raymond a très bien roulé en 4'03 et Anthony est  
 R. have.PRS.3SG very well drive.PTCP in 4'03 and A. be.PRS.3SG  
 descendu en 4'08.  
 go.down.PTCP in 4'08.

‘Raymond drove very well in 4’03 and Anthony est descendu en 4’08.’

In (77), the aspectual complement of duration in 4’08 impedes the agentive interpretation of the subject: it is unlikely for the subject to precisely determine the duration required to cover a given distance.

Expectedly, internal cause subjects associated with the spatial motion meaning of *descendre* tend to opt for ‘be’ (79). Those combined with *descendre* indicating ‘weight loss’ prefer rather ‘have’ (78), though not excluding ‘be’.

- (78) Le pilote était descendu pour être au dessous des nuages.  
 DEF.ART pilot be.PRS.3SG go down.PTCP to be.INF under.of DEF.ART clouds

‘The pilot has descended to below the clouds.’

- (79) Depuis ce matin, j’ai repris le régime [...] Il y  
 since this morning I.have- PRS.1SG restart.PTCP DEF.ART diet 3SG LOC  
 a trois semaines, j’ai descendu à 101.8 [...]  
 have.PRS.3SG three weeks I.have-.PRS.1SG go.down.PTCP to 101.8

‘Since this morning, I’ve restarted the diet [...] Three weeks ago, I went down to 101.8 [...].’

As to agentive subjects, there are 20 sentences with ‘have’ and 19 with ‘be’. The sentences with ‘be’ all represent spatial motion constructions.

The sentences with ‘have’ consist of (i) pseudo-transitive constructions with an implicit object associated to a quantity adverbial, e.g. *descendre (la résolution) à 5mpx* ‘lower (the resolution) to 5mpx’, and (ii) intentional actions, as illustrated in (80):

- (80) [...] le free secure [...] permis de faire de plus belles photos  
 DEF.ART free secure allow.PTCP to make of more beautiful photos
- en compressant moins les images (j'ai descendu à 5mpx).  
 compressing less DEF.ART images I.have-PRS.1SG go.down.PTCP to 5mpx
- '[...] the free\_secure [...]made it possible to take better photos with less image compression (I went down to 5 Mpx).'

Although the proportion between 'have' and 'be' is balanced in the sample of 100 sentences for agentive subjects, it is important to note that the auxiliary 'be' is largely dominant in the corpus, with 15,000 tokens compared to 2,000 with 'have' (including the transitive form). This suggests that the verb *descendre* 'go down' tends to favor 'be', like its Italian counterpart *scendere*.

However, instances of 'have' may occur for several reasons, even with human subjects, especially when they lack agentivity: firstly, 'have' can be used for *descendre* in a quantitative evolution construction, especially referring to weight loss or to other measurable evolutions. Secondly, many instances of 'have' in French can be attributed to the verb being used in a transitive-like construction with an internal object or an implicit object.

Furthermore, as observed previously, 'have' is an auxiliary is more grammaticalized in French. Lastly, the choice of auxiliary may also be influenced by the meaning conveyed, particularly for non-human subjects, featuring an increase in the use of 'have' with the figurative-derived meaning.

**(Lack of) Agentivity with non-human subjects** Human subjects exhibit a notable correlation with internal cause and agentive features. Conversely, non-human subjects are predominantly associated with a lack of agentivity and mostly occur with the auxiliary 'have', as Figure 4.16 illustrates.

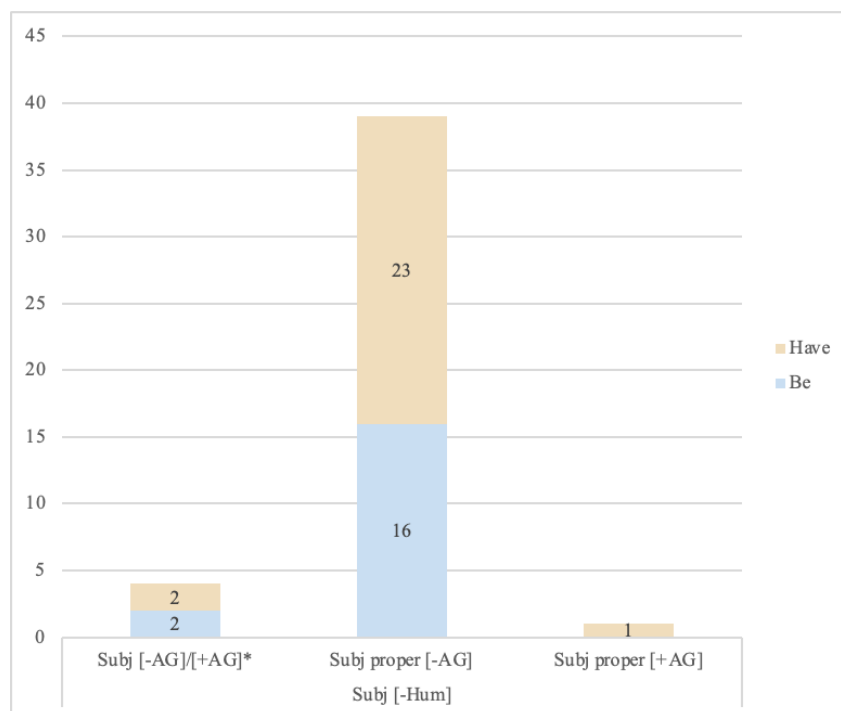


FIGURE 4.16 – Correlation of have and be, internal cause and non-human subjects.

Factors	Data 'Be'	Data 'Have'	Total
<b>Subj [-Hum] [-AG]/[+AG]*</b>	2	2	4
<b>Subj [-Hum] [proper -AG]</b>	16	23	39
<b>Subj [-Hum] [proper +AG]</b>	/	1	1
<b>Total</b>	18	26	44

TABLE 4.19 – Total occurrences relative to 'have' and 'be' according to internal cause, agentivity and human nature of the subject.

Table 4.19 demonstrates that nearly all instances with non-human subjects also feature non-agentive subjects. Although 'be' is possible, 'have' emerges as the preferred auxiliary. Examples (81) and (82) illustrate this phenomenon:

- (81) Ce n'est pas notre salaire qui a descendu, mais  
 it not.be.PRS.3SG NEG our wage that have.PRS.3SG go.down. PTCP but
- notre besoin de choses...  
 our need of things

'It's not our salary that has decreased, but our need for things...'

(82) Nous sommes concentrés, le stress est descendu on  
 we be,PRS.1PL focused DEF.ART stress be.PRS.3SG go.down. PTCP 3SG  
 y est!  
 LOC be.PRS.3SG

‘We’re focused, the stress is down, we’re in!’

(81) and (82) feature similar subjects but select different auxiliaries: ‘have’ in (81) and ‘be’ in (82). The prevalence of ‘have’ in such instances is expected, given that all these sentences embody a quantitative evolution construction with a derived meaning, which tends to favor ‘have’. ‘Be’ is however possible, in particular with a resultative meaning (82). The only instance of a non-human agentive subject is referring to an animal (83):

(83) Le Serin des Canaries Serinus canaria a une  
 DEF.ART S. of. DEF.ART C. S. c. have.PRS.3SG INDF.ART  
 longue queue, un bec couleur chair [...]. Il a  
 long tail INDF.ART beak colour flesh it have.PRS.3SG  
 descendu tranquillement le long de la tige.  
 go.down.PTCP quietly along DEF.ART stem

‘The Canary Serin Serinus canaria has a longer tail, a flesh-colored beak [...] It has descended quietly along the stem.’

The sentence in (83) exemplifies the use of ‘have’ in a spatial motion construction. It includes a locative adverbial *le long de la tige* referring to the path of the spatial movement rather than on its destination. Moreover, the adverb *tranquillement* ‘quietly’ implies an ongoing process. The use of ‘have’ is therefore aspectually motivated.

In conclusion, considering its semantic nuances, the verb *descendre* typically aligns with ‘be’ as its preferred auxiliary. However, in cases involving constructions conveying a derived meaning of quantitative evolution with non-agentive subjects, the usage of ‘have’ predominates.

**Adverbials** Among the various adverbials, the most frequent one is the adverbial expressing quantity. In second place comes the locative adverbial.

**ADVERBIAL OF QUANTITY** The adverbial of quantity is mostly associated to quantitative evolution constructions and, hence, to non-human subjects undergoing a quantitative evolution,

along different dimensions (ratio, weight, format, height (e.g., *les terrains ont descendu de plusieurs mètres* – ‘the ground fell several meters’).

Figure 4.17 illustrates the frequency of occurrences of ‘have’ and ‘be’ with the adverbial of quantity.

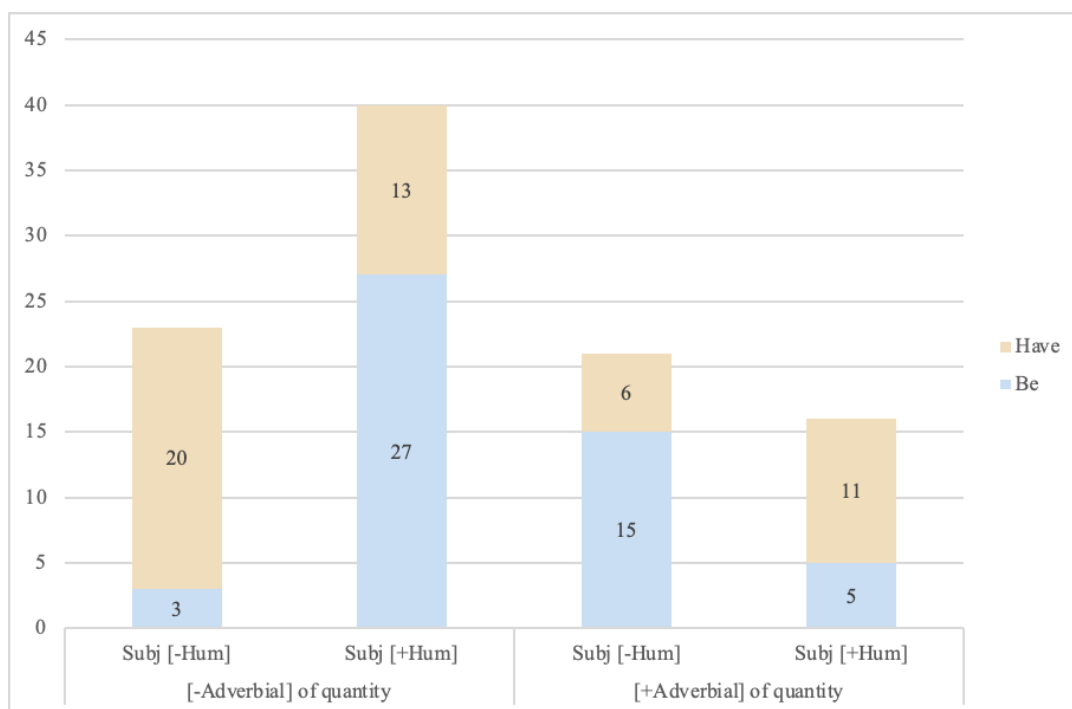


FIGURE 4.17 - Correlation of ‘have’ and ‘be’, the adverbial of quantity and the subject’s human trait.

Factors	Data ‘Be’	Data ‘Have’	Total
[-Adverbial of quantity] Subj [-Hum]	3	20	23
[-Adverbial of quantity] Subj [+Hum]	27	13	40
[+Adverbial of quantity] Subj [-Hum]	15	6	21
[+Adverbial of quantity] Subj [+Hum]	5	11	16
<b>Total</b>	50	50	100

TABLE 4.20 – Total occurrences relative to ‘have’ and ‘be’ according to the adverbial of quantity and the subject’s nature.

Surprisingly, there is a prevalence of ‘have’ with the adverbial of quantity and human subjects, while ‘be’ is more common with non-human subjects and the adverbial of quantity. This might at first sight seem contradictory to the patterns discussed earlier, but it can be explained.

First, a quantitative evolution construction does not require the presence of an adverbial of quantity, especially when the subject refers to a measure of quantity, e. g. *le taux de diabète a descendu* analyzed in previous sections, and its presence is, on the contrary, more crucial to

confer to *descendre* a quantitative meaning when the subject is not quantitative in nature, especially when the subject is human.

Secondly, both auxiliaries are possible with the quantitative evolution construction, as illustrated by (84) and (85), although ‘have’ is preferred because the meaning of quantitative evolution is a derived meaning.

(84) Depuis ma température est restée au alentour de 36.5 à  
 since.then my temperature be.PRS.3SG remain.PTCP.F.SG close.to 36.5 to

36.8 [...] Aujourd’hui elle a descendu à 36.4C.  
 36.8 today it have.PRS.3SG go.down.PTCP to 36.4C

‘Since then my temperature has remained around 36.5 to 36.8 [...] Today it’s down to 36.4C.’

(85) Le thermomètre est descendu à -20° près de Paris [...].  
 DEF.ART thermometer be.PRS.3SG go.down.PTCP to -20° near P.

‘The thermometer fell to -20° near Paris [...].’

The sentences in (84) and (85) exhibit similarities, showcasing the potential usage of both auxiliaries with non-human subjects. Despite the general preference of non-human subjects for ‘have’ (see Figure 4.14), a possible explanation for the presence of ‘be’ is that the construction in (85), involving a complement headed by the preposition *à*, closely resembles the spatial motion construction with a locative adverbial.

Another explanation is aspectual: the mention of an endpoint of the evolution yields a resultative state interpretation, which is compatible with the selection of ‘be’.

Considering human subjects, another reason accounting for the use of ‘have’ is that almost all occurrences combine with an internal object indicating quantity (86) or an implicit object (‘my medication’) associated with a complement indicating quantity (87):

(86=70) J’ai descendu 200m, puis remonté par le même  
 I.have.PRS.1SG descend.PTCP 200m then ascend.PTCP by DEF.ART same

itinéraire.  
 route

‘I descended 200 meters, then ascended by the same route.’

- (87) Au bout de trois jours à six capsules par jour j'ai eu  
 after three days at six capsules a day I.have.PRS.1SG have.PTCP
- des big crampe à l'estomac donc j'ai descendu  
 of.DEF.ART big cramps at DEF.ART.stomach so I.have.PRS.1SG go.down.PTCP
- à trois par jours [...].  
 to three a days

After three days at six capsules a day I got big cramps in my stomach, so I went down to three a day [...].

**LOCATIVE ADVERBIALS** Among locative adverbials, some evoke a static location, which can act as the arrival point for spatial motion (e.g., *près de Paris* 'near Paris'). whereas others are intrinsically directional (e.g., *jusqu'à la latitude de Terre-Neuve* 'up to the latitude of Newfoundland') or may even indicate a direction without specifying an endpoint (e.g., *on est descendus bas* 'we went down low').

In sentences with human subjects and locative adverbials indicating a static location, 4 out of 5 sentences select 'have', while 1 chooses 'be', contrary to the tendency where the presence of an endpoint correlates with 'be'. However, this preference can be understood if there are adverbials of manner that reinstate the focus on the process, as in (88).

- (88) Il a descendu pratiquement à 6km/h du sommet  
 He have.PRS.3SG go.down.PTCP in practice at 6km/h from.DEF.ART summit
- au Chalet Reynard.  
 to.DEF.ART C. R.

'He descended at almost 6km/h from the summit at the Chalet Reynard.'

This sentence contains two locative adverbials, *du sommet* 'from the summit' and *au Chalet Reynard* 'at the Chalet Reynard', the latter specifying the destination of the movement. All instances with a locative adverbial specifying an endpoint present either an adverbial of manner indicating the way the action has been performed (e.g., the speed of the downward movement) or an internal object. The data is too limited to support a hypothesis in terms of statistical trends, but it is plausible that an adverbial of manner indicating how the action has been performed highlight the ongoing process and hence favor 'have', even if the endpoint is indicated.

The sentence in (89) is another example of the locative adverbial specifying an endpoint, a manner adverbial of manner, selecting the auxiliary 'have':

(89) [...] j'ai descendu au Pertuiset à pieds.

I.have.PRS.1SG go.down.PTCP at.DEF.ART P. on.foot

'[...] I walked down to Le Pertuiset'.

Conversely, with a locative which does not involve an endpoint, 4 out of 5 occurrences with human subjects are with 'be' (90; 91):

(90) Les équipages de l'AG2R seront descendus aussi  
 DEF.ART crews of DEF.ART.AG2R be.FUT.3PL go.down.PTCP.M.PL so

bas pour rejoindre l'île de Saint-Barthélemy  
 low to reach.INF DEF.ART.island of S.B

'The AG2R crews will have gone this low to reach the island of Saint-Barthélemy'

(91) C'est vrai, on est descendu bas. Mais on a  
 It.be.PRS.3SG true we be.PRS.3SG go.down.PTCP low but we have.PRS.3SG

mieux rebondi que les socialistes.  
 better bounce.back.PTCP than DEF.ART socialists

'It's true, we've gone down. But we bounced back better than the Socialists.'

In (90), a spatial motion interpretation with a literal interpretation of the verb *descendre* is evident. Interestingly, (91) presents a spatial construction with a figurative interpretation, as previously analyzed. In any case, the spatial motion construction is associated with a primary meaning and tends to favor the auxiliary 'be'.

In summary, *descendre* exhibits patterns which are similar to other motion verbs analyzed thus far: in its spatial motion construction, the verb conveys a primary meaning, with a preference for the auxiliary 'be', particularly with human subjects.

In its quantitative evolution construction, typically involving non-human subjects and conveying a derived meaning, both 'have' and 'be' are possible, but 'have' is favored (table 4.21).

Construction	Human trait	Agentivity	Auxiliary
Spatial motion	Human	Agentive; Internal cause	Be
Quantitative evolution	Non-human	Non-agentive	Have

TABLE 4.21- Primary factors influencing the selection of 'have' and 'be' for *descendre*.

The table in 4.21 outlines the primary factors influencing the choice between ‘have’ and ‘be’, specifically focusing on the most frequently co-occurring conditions. Several factors converge to influence this choice: in cases of spatial motion constructions where the meaning is primarily locative, and the subject is human and displays agentive or internal cause characteristics, ‘be’ is the preferred auxiliary, although ‘have’ remains possible.

Conversely, in constructions related to quantitative change where the meaning typically is derived, and the subject tends to be non-human and non-agentive, both ‘have’ and ‘be’ are acceptable, with ‘have’ being the preferred choice.

Additionally, within the domain of quantitative evolution constructions, there exists the quantitative evolution construction with either a locative or a quantitative meaning, resulting in a pseudo-transitive structure, selecting ‘have’ (e.g., *J’ai descendu 200m*).

#### 4.1.5 Conclusion: Spatial Motion verbs

In this first section devoted to the auxiliary alternation in French (4.1), the analysis of the four spatial motion verbs reveals a major tendency among this category of verbs: the construction in which the verb is embedded emerges as the pivotal factor, thereby influencing other factors such as the nature of the subject.

The predominant constructions observed are those of spatial motion and quantitative evolution: spatial motion constructions predominantly employ the auxiliary ‘be’, whereas quantitative evolution constructions tend to favor ‘have’ (with the exception of *passer*).

Table 4.22 summarizes this tendency.

The primary influential factors for <i>Spatial Motion Verbs</i>	
CONSTRUCTION	AUXILIARY
Spatial Motion	‘be’
Quantitative Evolution	‘have’

TABLE 4.22 – Relevant factors for Spatial Motion Verbs

Undoubtedly, other factors highlighted during the analysis exert some influence, albeit minor, as they are contingent upon the type of construction. Thus, the type of construction stands as the primary determining factor, but the close link between construction types and the nature of the subject (in terms of  $\pm$ HUMAN,  $\pm$ AGENTIVE traits) indirectly results in a correlation between the choice of auxiliary and these parameters.

## 4.2 CHANGE-OF-STATE VERBS

Change-of-state verbs denote processes involving a transformation or alteration in the state or condition of an entity. Examples of such verbs in the corpus include *changer*, *mûrir*, *exploser*, and *diminuer*. With the exception of *exploser*, these verbs are classified as labile verbs, indicating that they can function both transitively and intransitively without altering their core meaning.

Depending on the construction, transitive or intransitive, the entity undergoing the change will be, from a syntactic point of view, either object or subject.

The present study is limited to the intransitive construction of these verbs. Unlike spatial motion verbs, change-of-state verbs have not developed multiple constructions with distinct meanings and differing in terms of auxiliary selection.

### 4.2.1 *Changer*

The verb *changer* ‘change’ demonstrates a notably higher frequency of occurrences with the auxiliary ‘have’ compared to ‘be’ in the FrTenTen (17) corpus. Upon applying a filter context of ‘no noun within five tokens right’, occurrences with ‘have’ are 91% (19,182 occurrences) compared to the 9% with ‘be’ (939 occurrences).

***Human and non-human subjects*** While ‘have’ dominates overall, the 100-sentences sample exhibits a relatively balanced distribution of auxiliaries depending on the human or non-human nature of the subject. Non-human subjects are more frequent and cooccur both with ‘have’ and ‘be’ auxiliated forms of *changer*, with a higher prevalence of ‘be’ (40 with ‘be’ and 34 with ‘have’), while for human subjects a slightly higher percentage of ‘have’ as an auxiliary (10 with ‘be’ and 13 with ‘have’) is observed.

A closer look at the data revealed that, with the verb *changer*, there are tokens that could be interpreted as transitive due to an implicit direct object. These tokens have been taken into consideration in order to analyze the subtle distinctions between the transitive and intransitive use of *changer*.

Even after excluding these possible transitive occurrences, the verb *changer* still exhibits a preference for ‘have’, as shown in Figure 4.18.

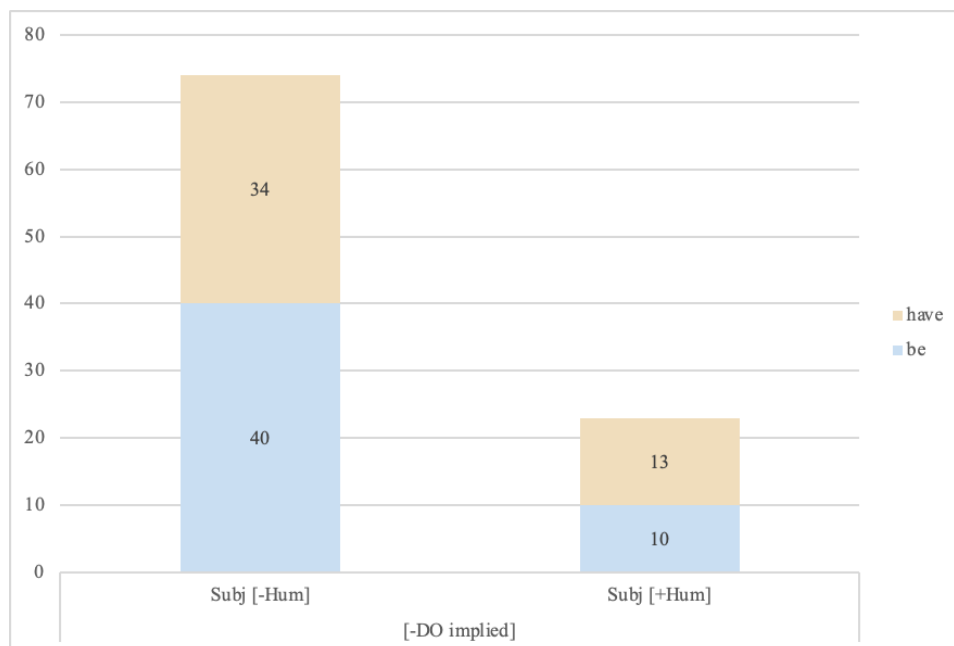


FIGURE 4.18 - Correlation of have and be, and the human/non-human nature of the subject.

Factors	Data 'Be'	Data 'Have'	Total
<b>[-DO] Subj [-Hum]</b>	40	34	74
<b>[-DO] Subj [+Hum]</b>	10	13	23
<b>Total</b>	50	47	97

TABLE 4.23 – Total occurrences relative to ‘have’ and ‘be’ according to the subject’s human nature.

Table 4.23 presents 13 occurrences with ‘have’ and 10 with ‘be’ with human subjects, compared to 34 with ‘have’ and 40 with ‘be’ for non-human subjects. The prevalence of ‘have’ results in more alternation between both auxiliaries in French compared to the Italian counterpart. Examples (92) and (93) illustrate instances of ‘have’ and ‘be’ with non-human subjects, respectively.

(92) À partir de ce moment-là tout a changé  
 From of this moment everything have.PRS.3SG change.PTCP

‘From that moment on, everything changed.’

(93) Tout est changé et surprend ici: des formats  
 everything be.PRS.3SG change.PTCP and surprise.PRS.3SG here some formats

carrés, des photographies de paysages qui n’existent pas.  
 squared of.DEF.ART photos of landscapes that not-exist.PRS.3PL NEG

‘Everything has changed and is surprising here: square formats, photos of landscapes that don’t exist’

Sentences (92) and (93) illustrate two sentences with the same subject (*tout* ‘everything’) and the use of the auxiliary verbs ‘have’ and ‘be’, respectively. The *passé composé* focuses on the process in (92), an interpretation triggered by the presence of a temporal complement referring to the past, and on the resulting state in (93), as witnessed by its coordination with the present tense. Hence, while there is an aspectual difference arising from the alternation between ‘have’ and ‘be’, there is no precise parameter that consistently determines the selection of one auxiliary over the other.

This observation similarly applies to sentences featuring a human subject, as demonstrated in (94) and (95). The sentence in (95) is particularly interesting as it features both auxiliaries, with the second person and the first-person singular respectively.

- (94) [...] tu as changé, avant on pouvait tout te  
 tu have.PRS.2SG change.PTCP before 3SG can.IPFV.3SG everything to.you
- dire et maintenant non [...]  
 tell.INF and now not

‘[...]you’ve changed, before we could tell you everything and now we can’t [...].’

- (95) Toi aussi, tu as changé? -Oui, je suis changée.  
 You too you have.PRS.2SG change.PTCP yes I be.PRS.1SG change.PTCP.F.SG

‘Have you changed too? Yes, I’ve changed.’

The fact that ‘have’ is more grammaticalized as a component of the *passé composé* in French has undoubtedly reduced the likelihood of encountering both auxiliaries in similar contexts, particularly in ‘change-of-state’ verbs like *changer*.

**Agentivity and internal cause** While agentivity and internal cause may still play a role, their influence is less pronounced compared to Italian due to the widespread usage of ‘have’ in French.

However, the agentivity factor does play a role: when the subject is non-agentive, ‘be’ is more frequent, while ‘have’ is favored with internal cause subjects, as shown in Figure 4.19 and Table 4.24.

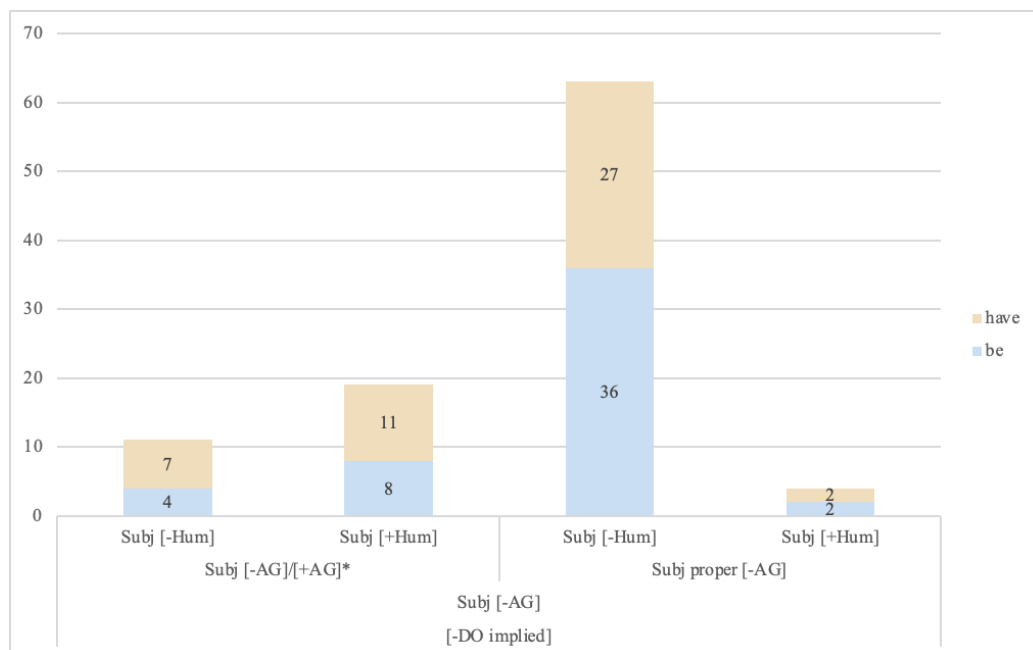


FIGURE 4.19 – Correlation of ‘have’ and ‘be’ with the subject’s nature (internal cause or non-agentive).

Factors	Data ‘Be’	Data ‘Have’	Total
[-DO] Subj [-AG] [-AG]/[+AG]*[-Hum]	4	7	11
[-DO] Subj [-AG] [-AG]/[+AG]*[+Hum]	8	11	19
[-DO] Subj [-AG] proper [-AG][-Hum]	36	27	63
[-DO] Subj [-AG] proper [-AG][+Hum]	2	2	4
<b>Total</b>	<b>50</b>	<b>47</b>	<b>97</b>

TABLE 4.24 – Distribution of ‘have’ and ‘be’ based on the subject’s nature (internal cause/not agentive and human/non-human) in intransitive uses.

The distribution of auxiliaries is characterized by the absence of agentivity. Although both ‘have’ and ‘be’ are present, there is a slightly higher percentage of ‘have’ when the subject is an internal cause (11 ‘have’ vs. 8 ‘be’ with human subjects; 7 ‘have’ vs. 4 ‘be’ with non-human subjects). Examples (96) and (97) illustrate this pattern:

(96) Nicolas Sarkozy a changé effectivement profondément.  
 N. S. have.PRS.3SG change.PTCP indeed profoundly

‘Nicolas Sarkozy has indeed changed profoundly.’

(97) Même s'il garde sa vieille peau, il est changé  
 Even if-he keep.PRS.3SG REFL old skin he be.PRS.3SG change.PTCP

intérieurement [...]   
 inside

‘Even if he keeps his old skin, he has changed inside [...].’

In (96) and (97), the human subject is categorized as an internal cause, as it could be interpreted as the instigator of the change. The adverbials *profondément* ‘deeply’ and *intérieurement* ‘internally’ reinforce this interpretation: the subjects can be both the entity creating the change and the entity impacted by it.

Similarly, non-human subjects categorized as internal causes also exhibit occurrences with both auxiliaries), albeit with a preference for ‘have’ (98). Examples (98) and (99) illustrate this distribution:

(98) Très cordial au cours d'un séminaire, son attitude a changé  
 Very cordial during. INDF.ART seminar REFL attitude have.PRS.3SG change.PTCP

du tout au tout [...].  
 completely

‘Very cordial during a seminar, his attitude changed completely [...].’

(99) [...] aujourd'hui son discours envers les femmes est  
 today REFL discourse towards DEF.ART women be.PRS.3SG

changé.  
 change.PTCP

‘[...] Today, his discourse towards women has changed.’

In (98) and (99), *attitude* ‘attitude’ and *discours* ‘discourse’ both refer to [+HUMAN] feature, and so they may act as internal causes. Changes in attitude and discourse inherently require human agency to bring about such alterations. In cases where the subject lacks agentivity, ‘be’ tends to predominate, especially with non-human subjects, although both auxiliaries may still occur.

Examples of non-agentive and non-human subjects with both auxiliaries are given in (100) and (101):

(100) [...] Même si sa vie a changé depuis.  
 even if his/her life have.PRS.3SG change.PTCP since.then

‘[...] even if his life has changed since then.’

(101) Notre vie est changée soudainement alors que nous nous  
 Our life be.PRS.3SG change.PTCP.F.SG suddenly when we REFL

y attendions le moins.  
 LOC expect.IPFV.1PL DEF.ART least

‘Our lives are suddenly changed when we least expect it.’

In (100) and (101), the subject *vie* ‘life’ is wholly non-agentive since life itself cannot actively initiate the change; it merely undergoes the change.

For comprehensive analysis, it is noteworthy that even human non-agentive subjects exhibit occurrences with both auxiliaries in this corpus, although the frequency is insufficient to draw definitive conclusions.

Examples (102) and (103) illustrate this point:

(102) [...] ce ne sont pas les électeurs qui ont changé,  
 it not be.PRS.3PL NEG DEF.ART voters who have.PRS.3PL change.PTCP

mais l’offre [...].  
 but DEF.ART.offer

‘[...] it’s not the voters who have changed, but the offer [...].’

(103) Je mets 4-0 à l’équipe[...] même le gardien est  
 I put.PRS.1SG 4-0 at.DEF.ART.team even DEF.ART goalkeeper be.PRS.3SG

changé[...].  
 change.PTCP

‘I put 4-0 to the team [...] even the goalkeeper is changed [...].’

In (102) and (103), although the subjects are human, they do not act as internal causes; rather, they are the entities undergoing replacement. The voters and the goalkeeper have changed because others have taken their place.

The non-agentive nature of the subjects allows for the analysis of the sentence in (105) as a passive construction. Since *changer* is a labile verb, sentences with non-agentive subjects and the auxiliary ‘be’ can indeed present a structural ambiguity between an intransitive and a passive construction.

In summary, the verb *changer* allows auxiliary alternation, demonstrated not only in this 100-sentences corpus but also in the larger FrTenTen (17) corpus. Both auxiliaries are observed with a significant level of frequency, but ‘have’ is more common.

In the 100-sentences sample, auxiliary selection is influenced by the nature of the subject. The prevalence of ‘have’ is particularly notable in cases involving internal cause subjects, while ‘be’ is more frequent with non-agentive subjects.

**Adverbials** The occurrences analyzed in this corpus do not exhibit many instances of adverbials. Among the few that appear, the most frequent ones are adverbials of manner and time adverbials.

**ADVERBIALS OF MANNER** The adverbial of manner, though infrequently, tends to favor ‘have’, as indicated in Figure 4.20 and table 4.25.

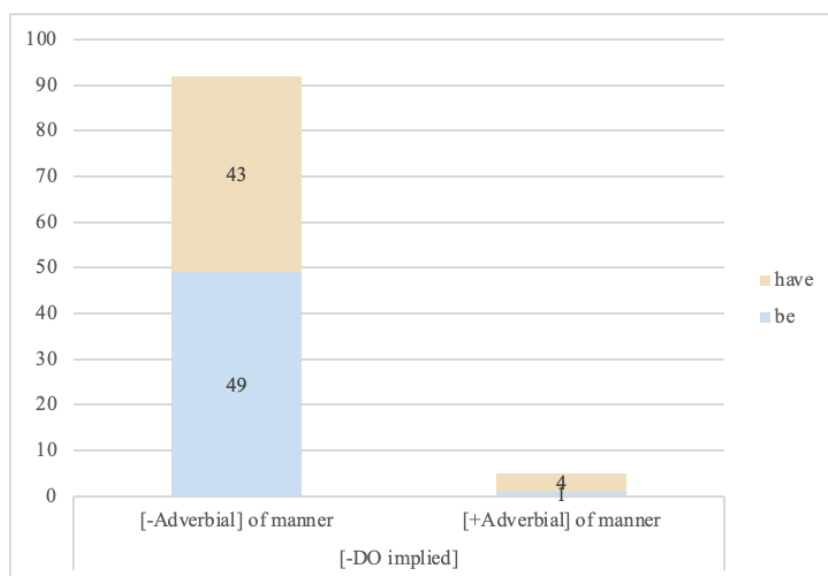


FIGURE 4.20 – Correlation of the ‘have’ and ‘be’ and the adverbial of manner.

Factors	Data ‘Be’	Data ‘Have’	Total
[-DO] - Adverbial of manner]	49	43	92
[-DO] + Adverbial of manner]	1	4	5
<b>Total</b>	50	47	97

TABLE 4.25 – Distribution of ‘have’ and ‘be’ based on the adverbial of manner.

Table 4.25 depicts the relationship between the adverbial of manner and the choice of auxiliaries. While there are only five instances of adverbials of manner in total, four out of five cases feature ‘have’, as exemplified in (104):

(104) Mais si nous pensons aux normes sociales[...] il devient  
 but if we think.PRS.1PL at.DEF.ART norms social 3SG become.PRS.3SG  
 clair qu'elles ont changé rapidement et radicalement.  
 clear that.they have.PRS.3PL change.PTCP rapidly and radically

‘But if we think about social norms [...] it becomes clear that they have changed rapidly and radically.’

Two factors appear to bias the choice towards ‘have’: the presence of an internal cause subject and the inclusion of an adverbial of manner, suggesting an ongoing process. This observation aligns with the hypothesis proposed for Italian verbs (seen for *baisser* as well), indicating that the interpretation of a process is more closely associated with ‘have’, while the resulting state tends to favor the use of ‘be’.

**TIME ADVERBIALS** The time adverbials appear in 13 instances of the sample data. When the subject is an internal cause, the preferred auxiliary is ‘have’, whereas ‘be’ is favored with non-agentive subjects (as well as in the absence of the time adverbial). Despite the contextualization of events by the time adverbial, the distribution of ‘have’ and ‘be’ remains similar both in its presence (7 occurrences with ‘be’ vs. 5 with ‘have’) and absence (43 occurrences with ‘be’ vs. 42 with ‘have’). Consequently, this adverbial is not considered as having an impact on auxiliary selection.

In conclusion, the verb *changer* in French demonstrates flexibility in its usage with both ‘have’ and ‘be’, although ‘have’ predominates, as indicated by the data from FrTenTen (17). The prevalence of ‘be’ is contingent upon two factors: non-agentive subjects, on the one hand, and passive voice and/or resultative aspect, on the other.

Table 4.26 summarizes these factors influencing the selection.

Construction	Human trait	Agentivity	Auxiliary
Ongoing process	Human/non-human	Non-agentive, Internal cause →	Have
Resultative/Passive	Human/non-human	Non agentive →	Be

TABLE 4.26 – Primary factors influencing the selection of ‘have’ and ‘be’

As stated in Table 4.26, the analysis of the 100-sentences corpus indicates that the nature of the subject can influence the choice of auxiliary. Specifically, a non-agentive nature tends to favor ‘be’.

Another crucial factor impacting auxiliary selection is aspect. Constructions perceived as representing an ongoing process tend to favor ‘have’, while those having a resultative meaning select ‘be’. A possible passive interpretation also favors ‘be’.

Hence the choice of ‘have’ or ‘be’ as an auxiliary seems to be more closely linked to the semantic interpretation of the construction rather than solely to the semantic features of the subject.

#### 4.2.2 *Mûrir*

*Mûrir* ‘ripen’/‘mature’ is a labile verb, meaning that it can be used both transitively and intransitively. However, its transitive construction is uncommon. As a change-of-state verb, it occurs more frequently with the auxiliary ‘have’ in FrTenTen (17). Specifically, ‘have’ auxiliation is strongly dominant, with a ratio of 96% occurrences compared to 4% occurrences with ‘be’ after the application of the part-of-speech filter. The occurrences found are 539 with ‘have’ and 21 with ‘be’.

To increase the dataset, all 80 occurrences of *mûrir* with the auxiliary ‘be’ found in the FrTenTen (17) corpus have been analyzed, since the issue of transitivity does not affect the use of ‘be’. However, only 40 of these occurrences met the criteria for inclusion in my study, resulting in a limited corpus of 90 occurrences for *mûrir*.

***Human and non-human subjects*** The human trait plays a significant role in auxiliary selection: while human subjects correlate exclusively with ‘have’, see example (105), non-human subjects exhibit variation, see (106) and (107), with ‘be’ being the more prevalent auxiliary, as Figure 4.21 illustrates.

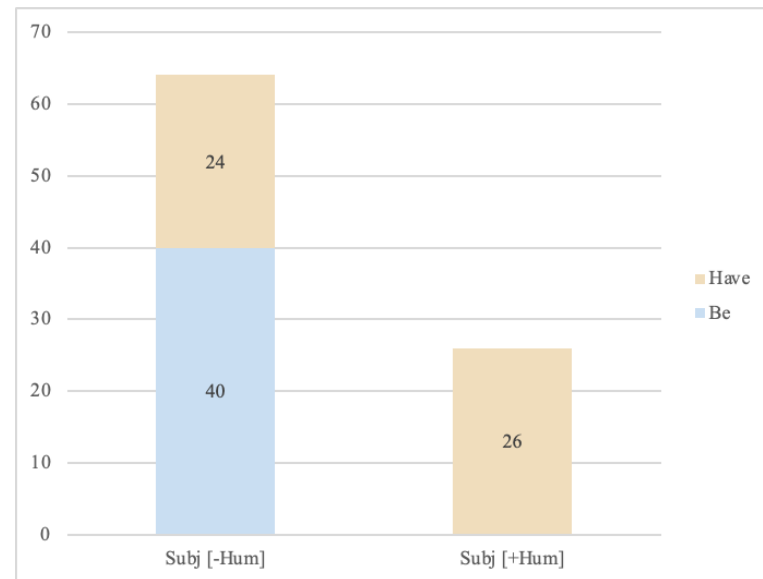


FIGURE 4.21 – Correlation of ‘have’ and ‘be’ with the human nature of the subject.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-Hum]</b>	40	24	64
<b>Subj [+Hum]</b>	/	26	26
<b>Total</b>	40	50	90

TABLE 4.27 – Distribution of ‘have’ and ‘be’ according to subject’s nature (human/non-human).

(105) [...] nous étions jeunes et immatures et nous avons  
 we be.IPFV.1PL young and immature and we have.PRS.1PL

mûri ensemble.  
 mature.PTCP together

‘[...] we were young and immatures, and we matured together.’

(106) Au cours d’un déjeuner, l’idée a mûri.  
 during. INDF.ART lunch DEF.ART.idea have.PRS.3SG mature.PTCP

‘During a lunch, the idea matured [...].’

(107) Les projets d’actions sont mûris collectivement avec  
 DEF.ART projects of.actions be.PRS.1PL mature.PTCP.M.PL collectively with

mûri habitants.  
 mature.PTCP residents

‘Action projects matured collectively with residents.’

The majority of non-human subjects associated with *mûrir* are characterized as ‘projects’ or ‘ideas’, entities that frequently have a strong link with a human referent. As depicted in Figure

6.21, *mûrir* predominantly combines with a subject corresponding to an internal cause, which, when non-human, is often closely associated with human activity.

**Agentivity, internal cause and lack of agentivity** Besides internal cause subjects, there are also instances with purely non-agentive subjects, but these are infrequent (only 3). It is reasonable to claim that the prototypical subject for *mûrir* corresponds to an internal cause.

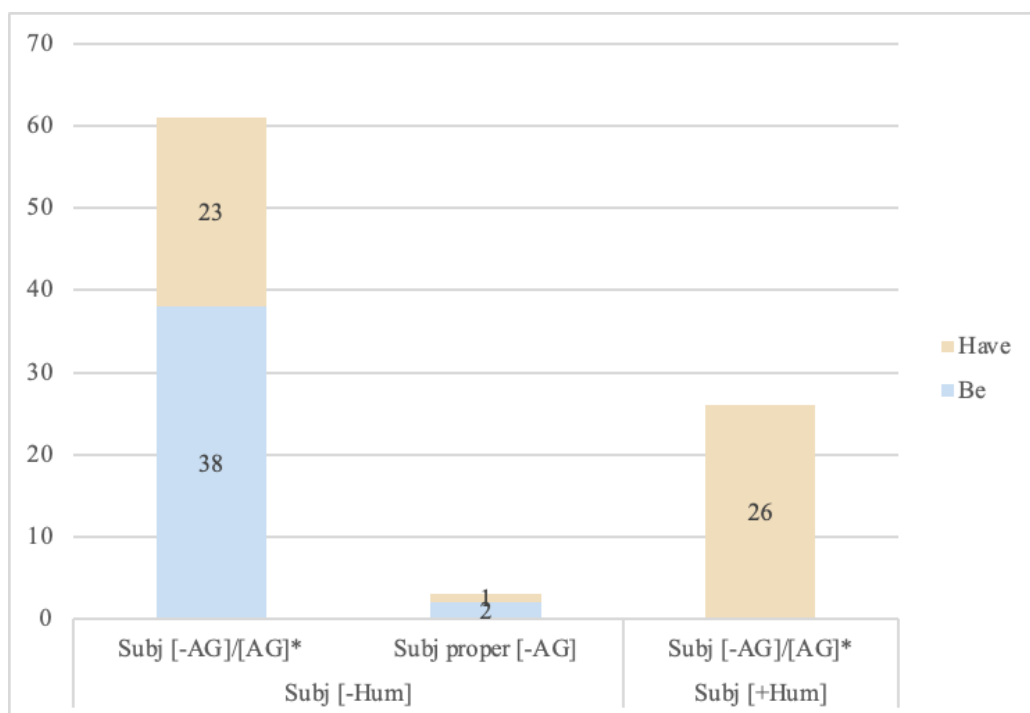


FIGURE 4.22 – Correlation of ‘have’ and ‘be’, internal cause and non-agentive subjects, and human/non-human subjects.

Factors	‘be’	‘have’	Total
Subj [-Hum] [-AG]/[+AG] *	38	23	61
Subj [-Hum] proper [-AG]	2	1	3
Subj [+Hum] [-AG]/[+AG] *	/	26	26
<b>Total</b>	40	50	90

TABLE 4.28 – Distribution of ‘have’ and ‘be’ according to the nature of the subject (human/non-human and agentive/internal cause)

When the subject is non-human and is an internal cause, the predominant auxiliary is ‘be,’ but the proportion of ‘have’ is significant, namely 38 occurrences with ‘be’ (108) and 23 with ‘have’ (109).

(108) L'idée a mûri en lisant quelques blogueuses.  
 DEF.ART.idea have.PRS.3SG mature.PTCP while reading a\_few bloggers

'The idea matured while reading a few bloggers.'

(109) Cela fait un peu plus de deux ans que ce projet  
 this do.PRS.3SG INDF.ART bit more of 2 years that this project

est mûri.  
 be.PRS.3SG mature.PTCP

'It has been two years since this project matured.'

In both (108) and (109), the subject corresponds to an internal cause: 'idea' and 'project' are strongly linked to the human referent conceiving them. An even more convincing example of an internal cause with the feature [+HUMAN] reference is provided in (110):

(110) Vingt-cinq ans déjà, au fil desquels on a l'impression  
 twenty-five years already through.which we have.PRS.3SG DEF.ART.impression

que le Parc a muri, qu'il a su  
 that DEF.ART park have.PRS.3SG mature.PTCP that.he have.PRS.3SG know.PTCP

parfaitement évoluer avec des attractions aussi diverses que  
 perfectly evolve.INF with DEF.ART attractions as diverse as

divertissantes.  
 entertaining

'Twenty-five years already, during which it feels like the Park has matured, evolving perfectly with attractions as diverse as they are entertaining.'

In (110), the subject *le Parc* is associated with the verb *savoir* 'know' 'be able', referring to an acquired ability and thus presupposing a human or at least animate subject. It is a clear example of attributing human actions to a non-human subject, e.g., *Le Parc*, considered as human because of the people behind it.

The tables and examples presented above indicate that 'have' and 'be' both occur with non-human subjects when the subjects are non-agentive but nevertheless internal causes, although 'be' tends to prevail. For human subjects, there is no variability, as the internal cause factor, combined with the human nature, can reinforce the control exerted by the subject.

Another factor playing a significant role is the aspect of the verb, as highlighted in academic literature, resultative aspect being associated to the 'be' auxiliary. Several of these factors (non-human subject, internal cause subject, and resultative aspect) may combine in one construction, and thus lead to the selection of the auxiliary 'be'.

An additional factor is illustrated in example (111), where an analysis as a passive construction, with an implicit agent ('the change has been matured by us'), rather than as an intransitive construction, is available.

- (111) Ce changement est mûri, car le débat s'est  
 this change be-PRS.3SG mature.PTCP as DEF-ART debate REFL.be.PRS.3SG  
 ouvert dès le plan de cohésion sociale.  
 open\_up.PTCP as\_soon\_as DEF.ART plan of cohesion social

'This change has been matured, as the debate opened up as soon as the social cohesion plan was launched.'

**Adverbials** The 90-sentences corpus of the verb *mûrir* contains few adverbials, including adverbials of manner, locative adverbials, and aspectual adverbials with progressive meaning.

**ADVERBIAL OF MANNER** Adverbials of manner, insofar as they manner highlights the verbal process rather than its result are expected to cooccur with 'have' rather than 'be'. The corpus contains only eight occurrences, with four instances featuring 'have' and three with 'be' when the subject is non-human.

Given the small sample size, it is not possible to infer tendencies, especially since other factors are at play in these eight occurrences.

As a matter of fact, auxiliary alternation is possible when the subject is non-human and internal cause. Human subjects cooccur invariably with 'have', whereas purely non-agentive subjects are barely attested.

Moreover, the variety of adverbials has to be taken into account. For instance, 'have' occurs especially with a gerund (112) or with the adverb suffixed with *-ment*, such as *imperceptibly* in (113), since both yield a focus on the process rather than the result.

- (112=108) L'idée a mûri en lisant quelques blogueuses.  
 DEF.ART.idea have-PRS.3SG mature.PTCP while reading a.few bloggers

'The idea matured while reading a few bloggers [...].'

- (113) Une succession/articulation entre deux projets qui a mûri  
 INDF.ART succession/articulation between two projects that have.PRS.3SG mature.PTCP  
 imperceptiblement [...].  
 imperceptibly

'A succession/articulation between two projects that imperceptibly matured [...].'

Other adverbials are compatible both with processive and resultative aspect (e.g., *collectivement* ‘collectively’ in (114), *avec BpiFrance* ‘with BpiFrance’), or may rather favor a resultative state interpretation (*très en amont* ‘beforehand’ in (115)).

- (114=107) Les projets d’actions sont mûris collectivement  
DEF.ART projects of.actions be.PRS.1PL mature.PTCP.M.PL collectively  
 avec les habitants.  
with DEF.ART residents  
 ‘Action projects matured collectively with residents.’

- (115) On y trouve d’autres valeurs: le projet étant mûri  
3SG LOC find.PRS.3SG of-others values DEF.ART project being mature.PTCP  
 très en amont avec Bpifrance [...].  
very beforehand with B.  
 ‘Other values are to be found here: the project was developed at a very early stage with Bpifrance [...].’

In these cases, the choice of auxiliary influences interpretation, leading us towards either a process interpretation (with ‘have’) or a resultative interpretation (with ‘be’).

**LOCATIVE ADVERBIALS** Locative adverbial may refer to a physical place (*à Paris*) or a metaphorical place (*dans votre cœur* ‘in your heart’). Examples (116) and (117) present the adverbial conveying a figurative meaning, while example (118) illustrates a literal meaning:

- (116) [...] l’histoire qui avait mûri en moi.  
DEF.ART.story that have.IPFV.3SG mature.PTCP inside me  
 ‘[...] the story that had matured inside me.’

- (117) [...] tout ce qui est mûri dans notre cœur à partir de  
everything what have.PRS.3SG mature.PTCP in our heart since

Campos  
 C.

‘[...] everything that has matured in our hearts since Campos.’

- (118) [...] les chambres où sont mûris des fruits.  
DEF.ART rooms where be.PRS.3PL ripen.PTCP.M.PL of DEF.ART fruit  
 ‘[...] rooms where fruit is ripened.’

In the sample of 13 occurrences containing a locative adverbial, both auxiliaries occur (10 occurrences with ‘be’ and 3 with ‘have’). Again, auxiliary alternation is restricted to non-human subjects mainly corresponding to an internal cause, e.g. (116; 117). (118) shows that ‘be’ yields structural ambiguity between the intransitive construction and the passive construction.

**ASPECTUAL ADVERBIAL WITH PROGRESSIVE MEANING** Adverbs expressing progression of the verbal process *progressivement* ‘progressively, *petit à petit* ‘little by little’ are expected to go hand in hand with the progressive aspect and therefore the auxiliary ‘have’. However, among the three occurrences, the auxiliary is ‘have’ in two of them (e.g. (119)) and (120).

(119) Et puis peu à peu j’ai mûri [...].  
and then little by little I.have.PRS.1SG mature.PTCP

‘And then little by little I matured [...].’

(120) Cette acceptation ensemble de la mort est mûrie  
This acceptance joint of DEF.ART death be.PRS.3SG mature.PTCP.F.SG

*progressivement*.  
gradually

‘This joint acceptance of death is a gradual process.’

Adverbs like *en un instant* ‘instantly’, *vite* ‘quickly’, *rapidement* ‘rapidly’ are expected to focus on the result of the verbal process, and to cooccur therefore with ‘be’.

Recall however that auxiliary alternation occurs only in combination with non-human subjects, whereas with a human subject, the auxiliary is invariably ‘have’, even if there is resultative aspect (121):

(121) Il voulait toujours aller avec les grands et, moi [...]  
he want,IPFV.3SG always go.INF with DEF-ART grown-ups and I

j’essayais toujours de l’emmener... Il a mûri très  
I.try.IPFV.1SG always to it.take.along-INF he have.PRS.3SG mature.PTCP very

*vite, bien trop vite*.  
quickly much too quickly

‘He always wanted to go with the grown-ups and I [...] always tried to take him along... He matured very quickly, much too quickly.’

To conclude, as in the case of *changer* and in contrast with the spatial motion verbs, the major factor influencing auxiliary selection is the nature of the subject. Human subjects, though infrequent, select invariably ‘have’, while non-human subjects allow both auxiliaries. In the latter case, the opposition between processive vs resultative aspect plays a major role, as well as the availability of an analysis as a passive construction.

Processive or resultative aspect is often not explicitly marked and may therefore be difficult to determine.

Table 4.29 summarizes the principal factors influencing auxiliary selection:

Construction	Human trait	Agentivity	Auxiliary
Ongoing process/resultative	Human	Internal cause →	Have
Ongoing process	Non-human	Internal cause →	Have
Resultative	Non-human	Internal cause →	Be

TABLE 4.29 – Primary factors co-occurring in the influence of the auxiliary for *mûrir*.

As shown on the basis of 100-sentences sample, when the subject is human, it consistently acts as an internal cause, and the verb exclusively selects ‘have’, irrespective of the opposition between processive or resultative aspect. Possible auxiliary alternation between ‘have’ and ‘be’ is associated with non-human, internal cause subjects.

Aspect proves to be a relevant factor. Focus on the ongoing process meaning tends to favor ‘have’, whereas focus on the resultative state favors ‘be’.

#### 4.2.3 *Diminuer*

The verb *diminuer* ‘decrease’, like *changer*, is a labile verb, meaning it can be used both transitively and intransitively. In this study, the focus is on its intransitive construction. Using the CQL and applying the part-of-speech filter ‘no noun in the first five tokens to the right’, the obtained sample shows that *diminuer* appears with the auxiliary ‘have’ in 83% of occurrences, while 17% of occurrences use ‘be’. Specifically, there are 6,069 instances with ‘have’ and 1,222 instances with ‘be’.

*Diminuer* combined with a subject referring to a measurable quantity is a verb denoting quantitative evolution, and close to *baisser*. Remember that the verb *baisser* literally denotes spatial motion, but it is predominantly used with its derived meaning in the quantitative evolution construction.

*Diminuer* combines less frequently with human subject. In this use it is properly a change-of-state verb, meaning ‘lose strength’, ‘weaken’, ‘decline’. *Diminuer* is compatible with both ‘have’ and ‘be’.

- (122) Le nombre de ventes a diminué de 7%.  
DEF.ART number of sales have.PRS.3SG decrease.PTCP by 7%

‘The number of sales decreased by 7%.’

- (123) Le taux de mortalité à tout âge est diminué de 10%.  
DEF.ART rate of mortality at any age be.PRS.3SG decrease.PTCP by 10%

‘The mortality rate at any age is reduced by 10%.’

The examples in (122) and (123) illustrate the quantitative evolution construction, this quantitative evolution being specified by an adverbial expressing quantity. Note, however, that the presence of this adverbial is not obligatory. By its intrinsic meaning of quantitative evolution, *diminuer* requires a non-human subject.

**Human and non-human subjects** When employed as an intransitive verb, *diminuer* primarily selects non-human subjects, the majority of which are inanimate, with one exception involving an animal. In this respect, *diminuer* is similar to *baisser*.

Interestingly, this verb demonstrates significant alternation between auxiliaries with non-human subjects, while occurrences with human subjects, albeit few in number, tend to favor ‘be’. Again, *diminuer* behaves in this respect like *baisser*.

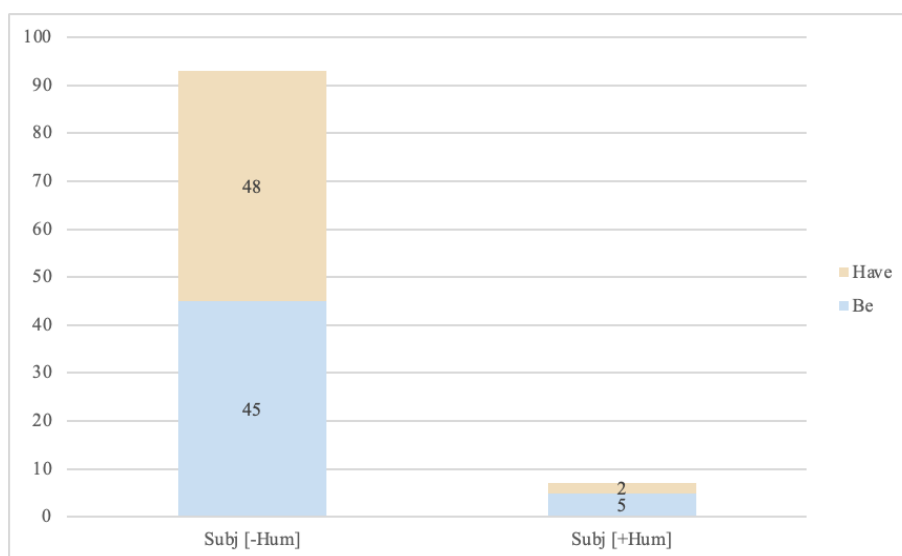


FIGURE 4.23 – Correlation of ‘have’ and ‘be’ with human and non-human subjects.

Factors	Data 'Be'	Data 'Have'	Total
Subj [-Hum]	45	48	93
Subj [+Hum]	5	2	7
Total	50	50	100

TABLE 4.30 – Distribution of ‘have’ and ‘be’ based on the subject’s human/non-human nature.

The use of ‘have’ or ‘be’ with human subjects triggers different readings: with ‘have’ (124), the meaning of ‘decrease’ remains intact, whereas with ‘be’, a resultative interpretation emerges, akin to an adjectival predication (125):

- (124) [...] le flot d’habitants fuyant vers le Sénégal a  
DEF.ART flow of.people fleeing to DEF.ART S. have.PRS.3SG

diminué.  
 decrease.PTCP

‘The flow of people fleeing to Senegal has reduced.’

- (125) [...] Il se rend bien compte qu’il est  
He REFL become.PRS.3SG well aware that.he be.PRS.3SG

diminué.  
 diminish.PTCP

‘[...] He is well aware that he is diminished.’

Examples (124) and (125) illustrate two distinct usages of the verb. In (124), the decrease in people fleeing to Senegal is depicted as a process, with the subject implied to be the number of people fleeing to Senegal. Conversely, in (125), the ‘be’ + past participle construction evokes a resultative state, conveying the meaning of ‘lose strength’.

Similar to *baisser*, *diminuer* is a labile verb, meaning it exhibits causative alternation and can function both transitively and intransitively. The intransitive form of such verbs can co-occur with either auxiliary in the *passé composé*, highlighting an aspectual contrast – one emphasizing an ongoing process and the other suggesting a completed result close to an adjectival predication.

Moreover, distinguishing whether the ‘be’ + past participle structure represents a resultative expression tied to the intransitive construction or is rather a passive form of the transitive construction can be challenging and indeterminate, as pointed out by Buchard & Carlier (2008:

2429). This distinctive feature becomes even more apparent when dealing with non-human subjects, as demonstrated in (126) and (127):

(126) [...] les coûts du sparging/ biosparging ont diminué  
DEF.ART costs of. DEF.ART sparging biosparging have.PRS.3PL decrease.PTCP  
 de 15%.  
 by 15%

‘[...] sparging / biosparging costs have decreased by 15%.’

(127) Les taux standardisés régionaux étaient diminués  
DEF.ART rates standardized regional be-IPFV.3PL decrease.PTCP.M.PL  
 d’au moins 20% [...]  
 by.at least 20%

‘Regional standardized rates decreased by at least 20% [...].’

Two observations can be made. Firstly, the intransitive construction with ‘be’ for *diminuer* resembles the passive construction. In (127), distinguishing between the intransitive use and the passive becomes challenging since the verb can be both transitive and intransitive. It is possible to assume the existence of an implicit external agent causing the decrease in rates in (127). Secondly, this ambiguity is absent in (126) due to the use of ‘have’. Therefore, auxiliary selection plays a pivotal role.

The situation could become more intricate when considering interpretations where internal cause subjects simultaneously act as a causer and undergo the action. While this has minimal implications for the semantics of the verb *diminuer*, it becomes more prominent in verbs like *changer* (in Section 4.2.1).

**(Lack of) Agentivity and internal cause** Considering that the verb is primarily employed in the construction of quantitative evolution or in a ‘be’ construction reminiscent of a passive structure, it is anticipated that the majority of subjects are non-agentive. This anticipation is consistent with the empirical findings, as depicted in Table 4.31.

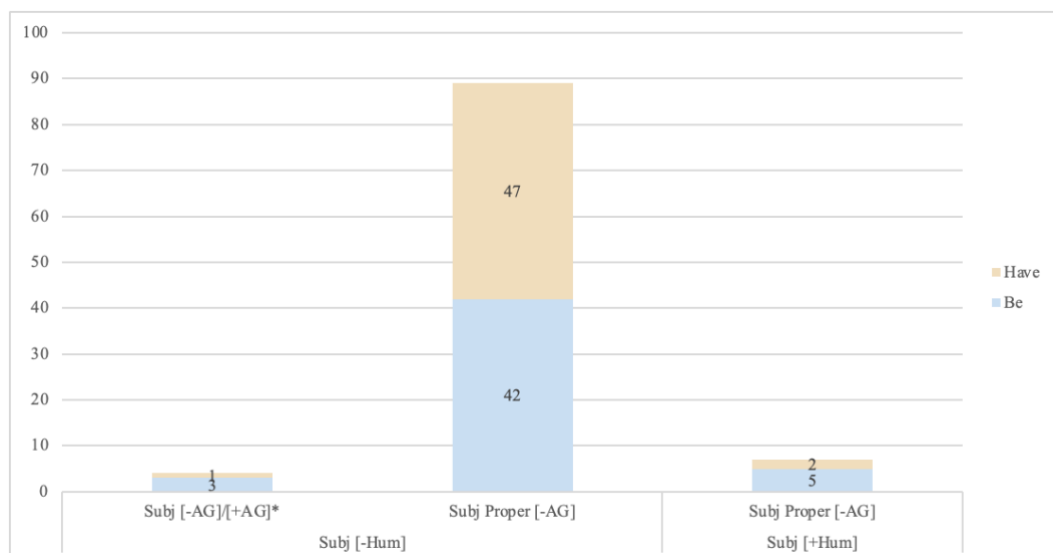


FIGURE 4.24 – Correlation of ‘have’ and ‘be’, non-agentive and internal cause subjects, human and non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-Hum] [-AG]/[+AG] *</b>	3	1	4
<b>Subj [+Hum] proper [-AG]</b>	42	47	89
<b>Subj [+Hum] proper [-AG]</b>	5	2	7
<b>Total</b>	50	50	100

TABLE 4.31 – Distribution of ‘have’ and ‘be’ according to the subject’s nature (human/non-human and internal cause/not agentive)

Nearly all instances exhibit non-agentive and non-human subjects. Non-human subjects with an internal cause interpretation are infrequent due to the verb’s semantics, which does not readily lend itself to a causative interpretation.

**Adverbials** In analyzing the most frequent adverbials, only non-human subjects were examined due to the scarcity of occurrences with human subjects. The predominant adverbial identified is the quantity adverbials.

**QUANTITY ADVERBIAL** The quantity adverbial is the primary adverbial that appears with the verb *diminuer*, given its association with quantitative evolution constructions, similar to what has been observed with *baisser*. As expected, the explicit mention of a quantity adverbial significantly increases the frequency of ‘have’. Figure 4.25 illustrates the prevalence of ‘have’ in the 59 instances featuring the adverbial of quantity.

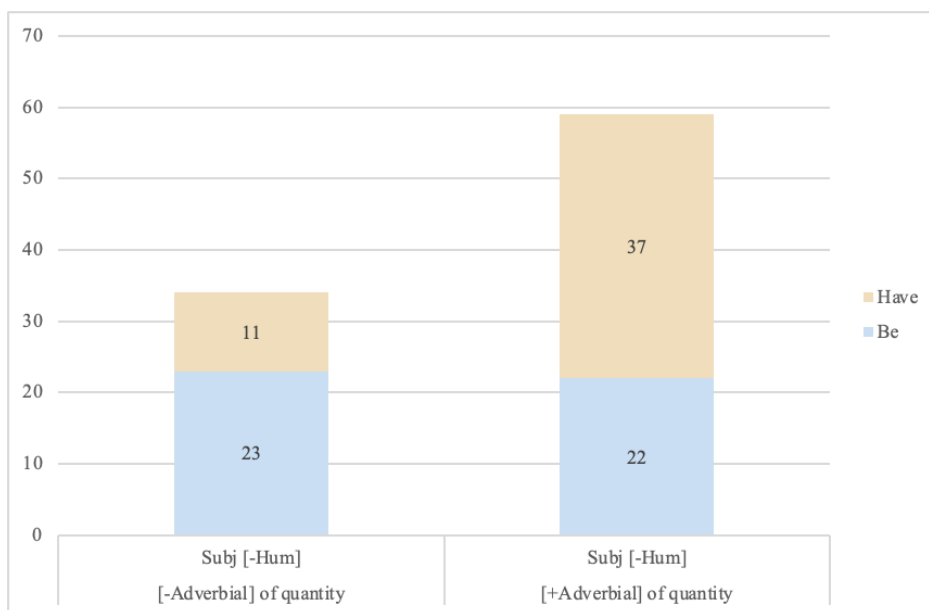


FIGURE 4.25 – Correlation of ‘have’ and ‘be’, the quantity adverbial and non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>[-Adverbial of quantity] Subj [-Hum]</b>	23	11	34
<b>[+Adverbial of quantity] Subj [-Hum]</b>	22	37	59
<b>Total</b>	45	48	93

TABLE 4.32 – Distribution of ‘have’ and ‘be’

Table 4.32 illustrates a preference for ‘have’ with 37 occurrences compared to 22 occurrences of ‘be’. However, it is worth noting that both ‘have’ and ‘be’ are possible, as demonstrated by examples (128) and (129).

- (128) En 2014, les dépenses nettes du FIVA ont  
 In 2014 DEF.ART expenses net of.DEF.ART F. have.PRS.3PL

diminué de 6%[...]  
 decrease.PTCP by 6% [...]

‘In 2014, FIVA’s net expenses decreased by 6% [...].’

- (129) Les taux de remboursement de soins [...] sont diminués  
 DEF.ART rates of reimbursement of care be.PRS.3PL decrease.PTCP.M.PL

de 40%.  
 by 40%.

‘Care reimbursement rates [...] decreased by 40%.’

Even when quantity is not expressed through adverbials, as in (130) and (131), the construction can still be quantitative. Comparing the two sentences clarifies how ambiguity fades when the auxiliary ‘have’ is used. In (131), the resulting state is highlighted.

- (130) [...]la distance que le son doit parcourir pour arriver  
DEF.ART distance that DEF.ART sound have.to.PRS.3SG travel.INF to reach.INF  
 jusqu’à vous a diminué.  
up.to you have.PRS.3SG decrease.PTCP  
 ‘[...] the distance sound has to travel to reach you has decreased.’

- (131) [...] la réponse au traitement est diminuée.  
DEF.ART response to. DEF.ART treatment be.PRS.3SG decrease.PTCP.F.SG  
 ‘[...] response to treatment is decreased.’

In summary, *diminuer* can be paired with both ‘have’ and ‘be’, especially correlating with non-human subjects.

The use of ‘be’ introduces constructional ambiguity, as the combination of ‘be’ with the past participle can be interpreted either as a resulting state linked to the intransitive construction or as a passive construction of the transitive form.

Table 4.33 illustrates the primary factors influencing the distribution of ‘have’ and ‘be’.

Construction or Interpretation <sup>23</sup>	Human trait	Agentivity	Auxiliary
Quantitative evolution	Human/Non-human	Non-agentive →	Have
Resultative/Passive	Human/Non-human	Non agentive →	Be

TABLE 4.33– Correlation of have and be with the adverbial of quantity and non-human subjects.

When the construction involves quantitative evolution, regardless of whether the subject is human or non-human, the subject is non-agentive, and the auxiliary can be both ‘have’ and ‘be’, although ‘have’ is preferred. However, when the construction is resultative or interpreted as passive, and the subject is non-agentive, the auxiliary used is ‘be’.

<sup>23</sup> Because it was necessary to present the correlations of the auxiliary in the same table, the terms construction and interpretation were used. Construction refers to quantitative and passive constructions, while interpretation refers to resultative.

#### 4.2.4 *Exploser*

The verb *explorer* ‘explode’ appears in 96% of occurrences with ‘have’ and 4% with ‘be’ in FrTenTen (17), using CQL and the additional filter context ‘no noun in 5 words right’. Specifically, there are 2,277 occurrences with ‘have’ and 101 occurrences with ‘be’.

Unlike other change-of-state verbs examined, *explorer* is intransitive without a corresponding transitive construction.

***Polysemy and different constructions*** For this verb, constructions such as quantitative evolution constructions have not been systematically analyzed, as it is not possible to make a real comparison with other types of construction, in contrast to the spatial motion verbs.

In general, *explorer* is associated with different meanings or constructions: the primary meaning (i), the derived or quantitative meaning/construction (ii), the semi-auxiliary use (iii), and the lexicalized use of *être explosé* (iv).

The primary or literal meaning is illustrated in (132) and (133). Both auxiliaries occur.

- (132) [...]ma petite voiture que j’ai pu me payer  
my little car which I.have.PRS.1SG can.PTCP REFL pay.INF  
  
grâce à la sueur de mon front est explosée [...].  
through. DEF.ART sweat of my brow be.PRS.3SG break.down.PTCP.F.SG

‘[...] my little car, which I was able to buy with the sweat of my brow, has broken down [...].’

- (133) [...] mais la bouteille a explosée et il y  
but DEF.ART bottle have.PRS.3SG explode.PTCP.F.SG and 3SG LOC  
  
en a eu partout dans l’habitacle.  
of.it have.PRS.3SG have.PTCP everywhere in DEF.ART.cabin

‘[...] but the bottle exploded, and there was some of it everywhere in the cabin.’

The second, derived meaning is represented by the quantitative evolution construction: the quantitative interpretation emerges either through the presence of a quantity adverbial (134) or because the subject corresponds to a quantifiable entity (135/136). Expectedly, the auxiliary ‘have’ is prevalent.

(134) En deux années, le chômage de longue durée frappant  
 In two years DEF.ART unemployment of long time striking

les moins de 25 ans a explosé de 72%.  
 DEF.ART less of 25 years have.PRS.3SG explode.PTCP by 72%

‘In two years, long-term unemployment among the under-25s has exploded by 72%.’

(135) Nicolas a baissé les impôts (au début) sans  
 N. have.PRS.3SG lower.PTCP DEF.ART taxes initially without

baisser les charges, et du coup la dette a  
 lower.INF DEF.ART expenses and as.a.result DEF.ART debt have.PRS.3SG

explosé.  
 skyrocket.PTCP

‘[...] Nicolas lowered taxes (initially) without reducing expenses, and as a result, the debt skyrocketed.’

(136) Certaines auberges proposent un lit à 20 €, multiplié par  
 some inns offer.PRS.3PL INDF.ART bed for 20 € multiply.PTCP by

4, le budget est explosé.  
 4 DEF.ART budget be,PRS.3SG blow.up-PTCP

‘Some inns offer a bed for €20, multiplied by 4, the budget is blown up.’

The third meaning, correlated with the type of subject (human), is the semi-auxiliary<sup>24</sup> use, mainly in the combination *exploser de rire* ‘to explode in laughter’ (137;138):

(137) [...] une des organisatrices a explosé de rire  
 INDF.ART of.DEF.ART organizers have.PRS.3SG explode.PTCP of laugh.INF

en me voyant si contente.  
 in me seeing so happy

‘[...] one of the organizers exploded with laughter when she saw me so happy.’

(138) Tout le monde sera explosé de rire tellement  
 Every DEF.ART world be.FUT.3SG explode.PTCP of laugh.INF so.much

vous la racontez mal!  
 you it tell.PRS.2PL badly

‘Everyone will be exploding with laughter at how badly you’re telling it!’

<sup>24</sup> In contrast to Italian, there is no dedicated chapter on semi-auxiliaries for French, as the selection between ‘have’ and ‘be’ in French is not influenced by the infinitive form of the verb.

In this use, ‘be’ has a high token frequency (10 uses of ‘be’ vs. 1 of ‘have’). This is also the case of the fourth use characterized by a process of lexicalization (139):

(139) Puis je suis explosé, lourd, entièrement vidé faute très  
 then I be-PRS.1SG blow.up.PTCP heavy completely empty.PTCP due very  
 certainement à mon alimentation du vélo.  
 certainly to my diet of. DEF.ART bike

‘Then I exploded, heavy, completely emptied, very likely due to my diet on the bike.’

*Explosé* is derived from the past participle of the verb *exploser* ‘explode’ and is used to emphasize the intensity of the physical and emotional experience. While *exploser* originates from the verb *exploser* ‘explode’, it seems to have developed a specific meaning, denoting profound fatigue or exhaustion, and it seems to have lexicalized into an adjective, whose relationship to the verbal process has become tenuous.

Figure 4.26 represents the auxiliary alternation related to different uses and constructions.

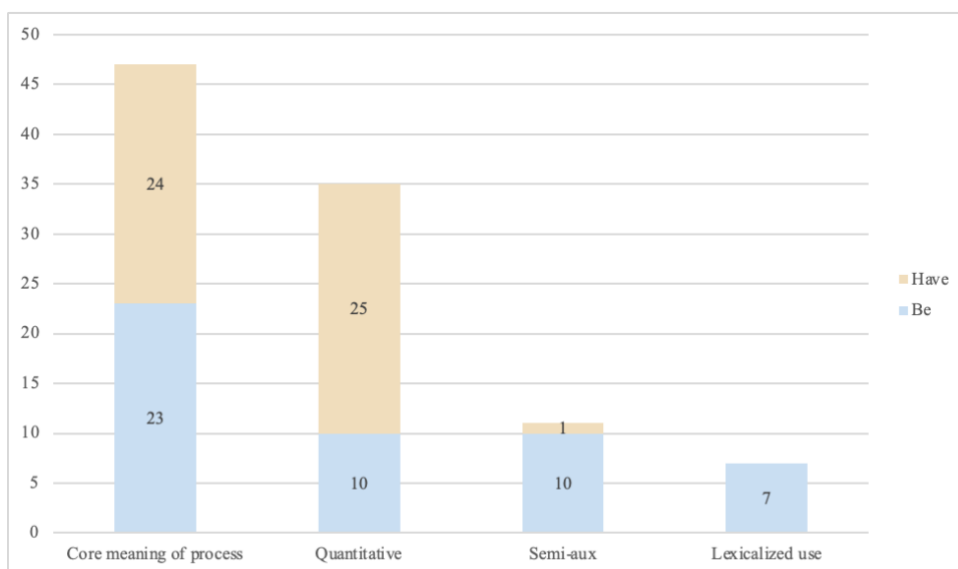


FIGURE 4.26 – Distribution of the auxiliaries based on the different constructions or meanings.

As mentioned earlier, the primary meaning of the verb represents the core meaning of the process.

**Human and non-human subjects** The frequency of *exploser* is higher in the sample with non-human subjects, where auxiliary alternation is most prominent. Globally, ‘have’ is more prevalent overall in FrTenTen (17). Figure 4.27 and Table 4.34 show that with a human subject,

‘be’ is nevertheless predominant, because the semi-auxiliary and lexicalized use occur only with human subjects. In contrast, with non-human subjects, ‘have’ prevails.

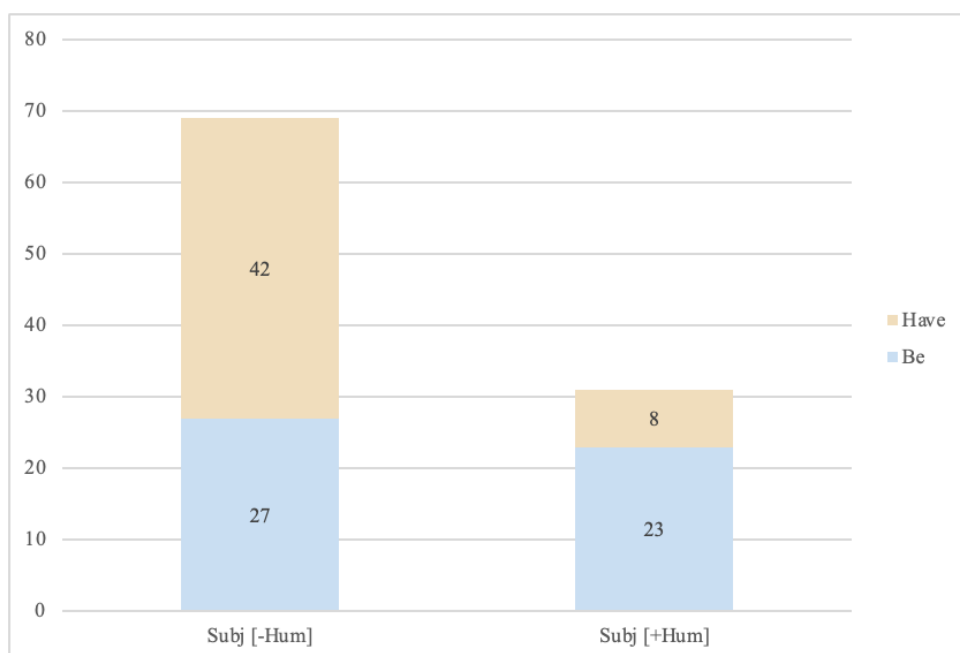


FIGURE 4.27 – Correlation of ‘have’ and ‘be’ with human and non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [-Hum]	27	42	69
Subj [+Hum]	23	8	31
<b>Total</b>	50	50	100

TABLE 4.34 – Distribution of ‘have’ and ‘be’ according to the subject’s human nature.

When the verb is used in its core process meaning, occurrences in the 100-sentence sample exhibit similar uses with ‘have’ (140) and ‘be’ (141) with human subjects:

(140) On a tous à l’esprit ces vies brisées, ces  
 3SG have.PRS.3SG all to DEF.ART -mind those lifes shattered those

familles qui ont explosé.  
 families who have.PRS.3PL explode.PTCP

‘We all remember those shattered lives, those families that exploded.’

(141) Mais chère Marie-Christine, le Front de Gauche est explosé.  
 But dear M.C. DEF.ART Front de Gauche be.PRS.3SG explode.PTCP

‘But dear Marie-Christine, the Front de Gauche is exploded’.

With non-human subjects, both ‘have’ and ‘be’ are used. The examples (142) and (143) showcase ‘have’ (142) and ‘be’ (143) with the primary meaning of *explorer*, whereas (144) and (145) feature both ‘have’ and ‘be’ with its derived meaning.

(142=133) [...] mais la bouteille a explosée et il y  
 but DEF.ART bottle have.PRS.3SG explode.PTCP.F.SG and 3SG LOC  
 en a eu partout dans l’habitacle.  
 of.it have.PRS.3SG have.PTCP everywhere in DEF.ART.cabin

‘[...] but the bottle exploded, and there was some of it everywhere in the cabin.’

(143) [...] les roches sont chaudes, un cratère est  
 DEF.ART rocks be.PRS.3PL hot INDF.ART crater be.PRS.3SG  
 explosé à 5 mètres de nous.  
 explode.PTCP at 5 meters from us

‘[...] the rocks are hot, a crater exploded 5 meters from us.’

(144) Essence, électricité, chauffage, les prix ont explosé.  
 gasoline electricity heating DEF.ART prices have.PRS.3PL explode.PTCP

‘Gasoline, electricity, heating: prices have exploded.’

(145) [...] au bout de 10 jours mon forfait est explosé.  
 after 10 days my package be.PRS.3SG explode.PTCP

‘[...] after 10 days my package is exploded!’

As the examples demonstrate, both auxiliaries are observed; however, ‘have’ is the preferred one.

**(Lack of) Agentivity and Internal cause** Figure 4.28 and Table 4.35 show that the majority of instances involving the verb *explorer* feature a non-agentive subject, viz. 90 instances, regardless of whether it is human or non-human. Instances of internal cause subjects are relatively infrequent, viz. 10 instances.

Figure 4.28 and Table 4.35 also show the frequency of both auxiliaries in relation to the nature of the subject, according to the features [ $\pm$ HUMAN], [ $\pm$ AGENTIVE].

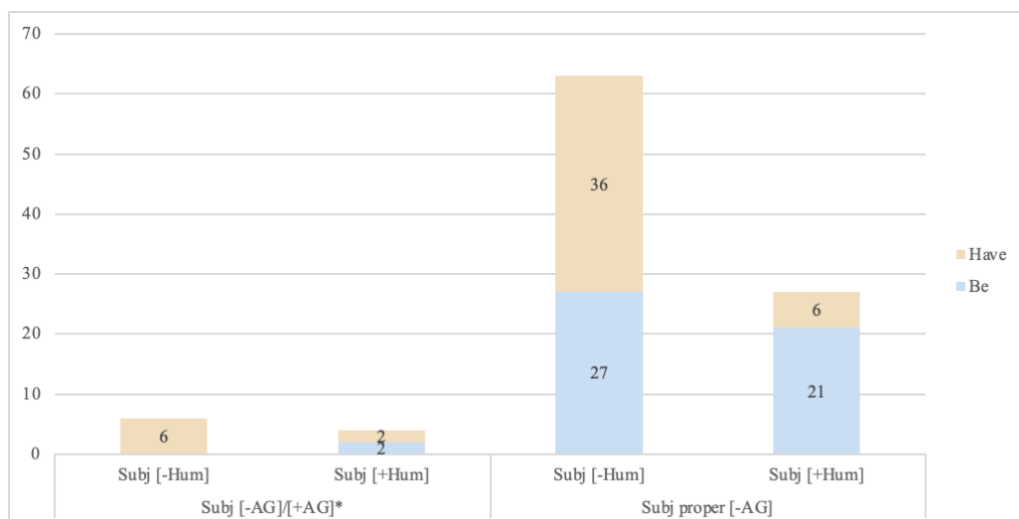


FIGURE 4.28 – Frequency of ‘have’ and ‘be’ according to [±HUMAN] and [±AGENTIVE].

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Subj [-AG]/[+AG] * [-Hum]</b>	-	6	6
<b>Subj [-AG]/[+AG] * [+Hum]</b>	2	2	4
<b>Subj proper [-AG] [-Hum]</b>	27	36	63
<b>Subj proper [-AG] [+Hum]</b>	21	6	27
<b>Total</b>	50	50	100

TABLE 4.35 – Frequency of ‘have’ and ‘be’ according to [±HUMAN] and [±AGENTIVE].

As to non-agentive subjects, ‘have’ prevails when the subject is non-human, whereas ‘be’ is dominant for human subject, in part because of the high token-frequency of two idiomatic uses of *exploser*, viz. *X[+Human] est explosé*; *X[+Human] est explosé de rire*.

When the observation is restricted to sentences with the verb’s core meaning, both human and non-human subjects present an equal number of ‘have’ and ‘be’ (18 occurrences of ‘have’ vs. 18 of ‘be’ for non-human subjects; 6 occurrences of ‘have’ vs. 6 of ‘be’ with human subjects).

Human and non-agentive subjects conveying the core meaning of process are illustrated with ‘have’ (146) and ‘be’ (147), respectively.

(146=140) On a tous à l’esprit ces vies brisées, ces  
 3SG have.PRS.3SG all at DEF.ART -mind those lifes break.PTCP.F.PL those  
 familles qui ont explosé.  
 families who have.PRS.3PL explode.PTCP

‘We all remember those broken lives, those families that exploded.’

(147) [...] c'est un symptôme quand même d'une société qui  
 it.be.PRS.3SG INDF.ART symptom anyway of.INDF.ART society that  
 est explosée.  
 be.PRS.3SG explode.PTCP.F.SG  
 '[...] anyway, it is a symptom of an exploded society.'

In (147), the subject is entirely non-agentive, as it is not primarily responsible for the action. In fact, the preceding expression, *ces vies brisées* 'these broken lives', suggests that the subject of *exploser* experienced something without actively controlling it.

A resultative interpretation is underscored in (147) with the auxiliary 'be', resembling (but not being) an adjectival predicate construction. This resultative interpretation is even more pronounced with non-human subjects in (148) and (149), which typically do not exert control over the verbal process even in the core meaning of process use.

(148=132) [...]ma petite voiture que j'ai pu me payer  
 my little car which I.have.PRS.1SG can.PTCP REFL pay.INF  
 grâce à la sueur de mon front est explosée [...].  
 through. DEF.ART sweat of my brow be.PRS.3SG break.down.PTCP.F.SG  
 '[...] my little car, which I was able to buy with the sweat of my brow, has broken down [...].'

(149) [...] il n'est pas dit que ce soit bien  
 3SG not.be.PRS.3SG NEG say.PTCP that this be.SUBJ.3SG well  
 une bombe nucléaire qui est explosée.  
 INDF.ART bomb nuclear that be.PRS.3SG explode.PTCP.F.SG  
 '[...] there's no guarantee that it was a nuclear bomb that exploded.'

In (148) and (149), the use of the 'be' auxiliary may accentuate resultativity. Considering the semantics of the verb *exploser* 'explode', 'have' may underscore the process, while 'be' emphasizes the resulting state. This differentiation is most pronounced when the past participle is used as an adjectival predicate, as evidenced in (148).

Internal cause subjects, actively participating in the process described by the verb, do occur in our sample with both 'have' and 'be' when the subject is human (151, 152), and with 'have' and when the subject is non-human (150), but the number of occurrences is too small to infer a solid hypothesis.

(150) L'utilisation du smartphone a explosé et contribue  
DEF.ART.use of,DEF.ART smartphone have-PRS.3SG explode-PTCP and help-PRS.3SG  
 aujourd'hui à modifier le parcours traditionnel d'achats.  
today to change-INF DEF.ART path traditional of.shopping

'The use of smartphones has exploded and is now helping to change the traditional shopping experience.'

(151) C'est lorsque son mari s'est exprimé bruyamment  
it.be,PRS.3SG when his husband REFL.be-PRS.3SG speak.PTCP loudly  
 qu'elle a explosé.  
that.she have,PRS.3SG explode.PTCP

'It was when her husband spoke loudly that she exploded.'

(152) [...] on s'enguelait comme un vieux couple devant toute  
3SG REFL.argue.IPFV.3SG like INDF.ART old couple in.front.of whole  
 la classe, la prof était explosée.  
DEF.ART class DEF.ART teacher be.IPFV.3SG explode.PTCP.F.SG

'[...] we were arguing like an old couple in front of the whole class, and the teacher exploded.'

In sentences (150), (151), and (152), the subjects are categorized as internal cause subjects, as they are both agents and patients of the action. Notably, in (150), the subject 'usage' implies a person using the smartphone, indicating a distinct reference to a [+HUMAN] trait.

**Adverbials** Examining the most frequent adverbials, only the time adverbials stand out. However, this category of adverbials does not influence auxiliary selection: human subjects with temporal adverbials tend to prefer 'be' (as well as in the absence of time adverbials), while non-human subjects favor 'have', regardless of any other conditions.

To conclude, *exploser* occurs more frequently with 'have', not only in the global FrTenTen (17) corpus, but also in the 100-sentences corpus, particularly with non-human subjects. With human subjects, however, the most frequent auxiliary is 'be', but this tendency is due to a high token-frequency of two idiomatic expressions.

Table 4.36 provides a summary of the factors co-occurring in the selection of ‘have’ and ‘be’.

Construction or Interpretation	Human trait	Agentivity	Auxiliary
Core meaning of process	Human/non-human	Not-agentive ➡	Have/Be
Lexicalized use ‘be’ + <i>explosé</i>	Human	Not-agentive ➡	Be
Semi-auxiliary use	Human	Not-agentive ➡	Be
Quantitative	Non-human	Not-agentive ➡	Have

TABLE 4.36- primary factors influencing the selection of ‘have’ and ‘be’

In Table 4.36, the primary and most frequent factors co-occurring with the use of the auxiliary are presented. The first crucial factor is the type of construction: *exploser* has been found to appear in constructions representing its core meaning of process, in the lexicalized use ‘be’ *explosé*, in the semi-auxiliary use, and in the quantitative evolution construction. The most interesting one for *exploser* as a change-of-state verb is the core meaning of process. For this construction, the subject is consistently non-agentive, whether human or non-human. When the subject conveys the core meaning of the process delineated by the verb *exploser*, both ‘have’ and ‘be’ appear as auxiliaries.

In the lexicalized use ‘be’ + *explosé* and the semi-auxiliary use, the subject is human and the auxiliary tends to be ‘be’. The frequency of ‘be’ is due to idiomatic uses: the use of *explosé* as a predicative adjective (e.g., *je suis explosé, entièrement vidé*) or as a semi-auxiliary combined with the infinitive verb *rire* (e.g., *je suis explosé de rire*).

Finally, with the quantitative evolution construction, the subject is non-human and ‘have’ is the favored auxiliary.

#### 4.2.5 Conclusion: Change-of-State verbs

The analysis conducted in this section 4.2 reveals that for *changer* and *mûrir*, the most influential factor determining the choice of auxiliary is the semantic feature of internal cause. In this regard, these verbs exhibit similarities to the Italian peripheral verbs analyzed in the previous chapter. Specifically, when the subject is an internal cause, the preferred auxiliary is ‘have.’ Conversely, when the subject lacks agency, the auxiliary selected is ‘be’ (table 4.37).

*Diminuer* presents in all respects a profile reminiscent of that of *baisser* and behaves like other spatial motion verbs when used in a quantitative evolution construction.

As to *exploser*, its behavior depends on the type of construction or meaning conveyed: in a quantitative evolution construction, the preferred auxiliary is ‘have’; when used in an idiomatic expression or in a semi-auxiliary use, it prefers ‘be’; and it presents both auxiliaries when conveying its core meaning of the change-of-state ‘explode’.

Most influential factor for <i>change-of-state</i> verbs	
FACTORS	AUXILIARY
Internal Cause ( <i>changer, mûrir</i> )	‘have’
Lack of agentivity	‘be’
(Aspect)	Processive aspect: ‘have’
	Resultative aspect: ‘be’

TABLE 4.37 – Most influential factor for *change-of-state* verbs.

While other factors, such as the [ $\pm$ HUMAN] nature of the subject are important, they are most often a consequence of the internal cause/lack of agentivity, as a subject with an internal cause is more likely to be human or have a strong link to a human referent.

Another residual factor, highlighted in the table since it is distinct from the semantic nature of the subject, is the aspect of the verb. A challenge for my analysis is the rarity of adverbials corroborating independently either a resultative aspect or a processive aspect in the corpus data.

However, the general tendency is that an ongoing process interpretation favors ‘have’ and a resultative one favors ‘be’.

### 4.3 VERBS OF APPEARANCE

*Paraître* ‘appear’/‘seem’ and *apparaître* ‘appear’/‘show up’ are both verbs of appearance, describing the way something seems or appears. They convey the meaning of perception, observation, or the outward manifestation of something. A notable feature of these verbs is their ability to be used in a copulative construction.

#### 4.3.1 *Paraître*

In FrTenTen (17), the verb *paraître* exhibits 83% of occurrences with ‘have’ compared to 17% of ‘be’ when using CQL and applying the part-of-speech filter, showing a predominance for ‘have’. Specifically, there are 10,214 occurrences with ‘have’ and 2,093 with ‘be’.

The verb *paraître* appears in two distinct constructions: a non-copulative and a copulative one.

**Non-copulative construction (+ adverbials)** The non-copulative construction, featuring the verb *paraître* often accompanied by a time adverbial, conveys the meaning of ‘appear’. This construction almost exclusively selects ‘be’, as illustrated by Figure 6.28.

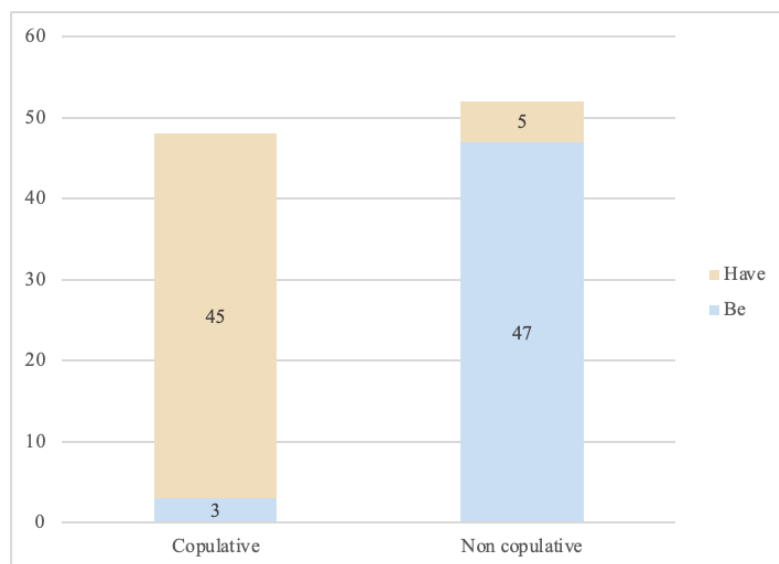


FIGURE 4.29 – *Paraître*: Frequency of ‘have’ and ‘be’ according to the construction’s type.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Copulative</b>	3	45	48
<b>Non copulative</b>	47	5	52
<b>Total</b>	50	50	100

TABLE 4.38 – *Paraître*: Frequency of ‘have’ and ‘be’ according to the construction’s type.

The non-copulative construction of the verb *paraître* is exclusively employed with non-human subjects, and is frequently accompanied by a time adverbial, as exemplified in (153):

- (153) Un           nouvel album du           groupe est           paru           en 2009.  
INDF.ART   new       album   of.DEF.ART   band   be.PRS.3SG   appear.PTCP   in   2009

‘A new band album appeared/was released in 2009’.

**Copulative construction (+dative)** The copulative construction serves to establish a syntactic linkage between the subject and a predicate. Within this construction, the semantic function attributed to the verb, in comparison with the standard copula ‘be’, is that of an evidential marker, close to ‘seem’. As shown in Figure 4.29 and Table 4.38, the copulative construction exhibits a clear predilection for the auxiliary ‘have’ (154).

(154) Elle a paru très surprise, mécontente [...]  
 She have.PRS.3SG appear.PTCP very surprised unhappy

‘She seemed very surprised, unhappy [...]’

Subjects within the copulative construction may be both human and non-human referents. In the majority of cases—though exceptions do exist—the construction incorporates a dative complement (155):

(155) Le débat [...] m’a paru complètement irréal.  
 DEF.ART debate [...] to.me.have.PRS.3SG seem.PTCP completely unreal

‘The debate [...] seemed completely unreal to me.’

**Human and non-human subjects** While non-human subjects appear in both copulative and non-copulative constructions, human subjects exclusively feature in the copulative construction. Notably, the attribution of human traits does not significantly influence the selection of auxiliaries. Figure 4.30 and Table 4.39 show the fact that it is the construction itself that dictates the choice between the auxiliary verbs ‘have’ and ‘be’.

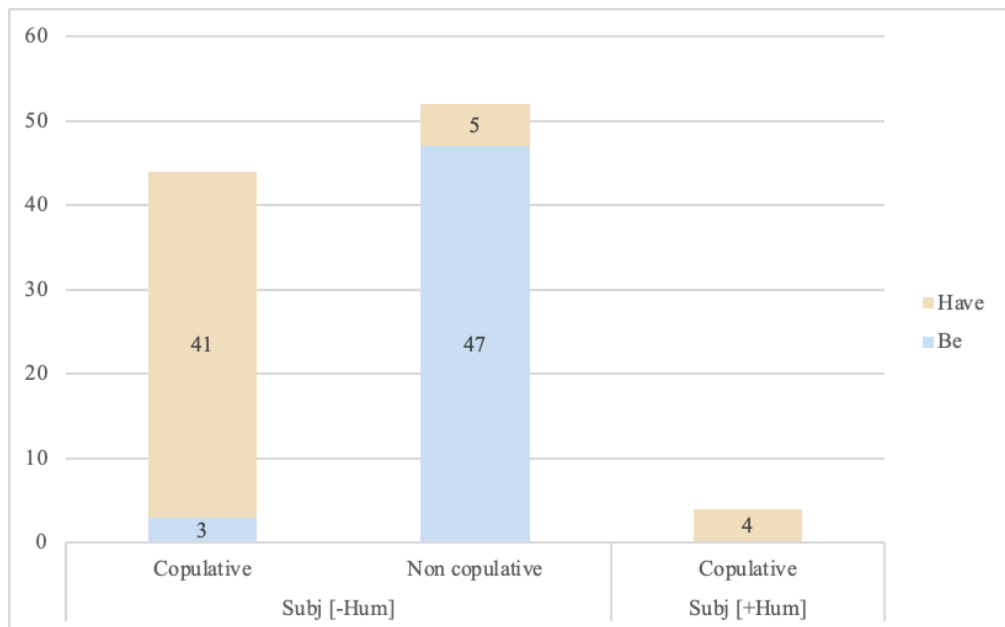


FIGURE 4.30 – *Paraître*: Frequency of ‘have’ and ‘be’, according to the construction’s type and the human trait of the subjects.

Factors	Data 'Be'	Data 'Have'	Total
Subj [-Hum] Copulative	3	41	44
Subj [-Hum] Non copulative	47	5	52
Subj [+Hum] Copulative	/	4	4
<b>Total</b>	50	50	100

TABLE 4.39 – *Paraître*: Frequency of ‘have’ and ‘be’, according to the construction’s type and the human trait of the subjects.

Subjects within the copulative construction exhibit a broad spectrum, encompassing not only human and non-human entities but includes impersonal *il*, as evidenced in examples (156) and (157).

- (156) Il nous a paru opportun de reprendre pour le  
3SG to.us have.PRS.3SG seem.PTCP appropriate to retake.INF for DEF.ART  
présent article le titre de la thèse de Bernard Lamblin.  
this article DEF.ART title of DEF.ART thesis of B. L.

‘It seemed appropriate to use the title of Bernard Lamblin’s thesis for this article.’

- (157) Toutefois, [...] il a paru possible de ne pas opposer  
however 3SG have.PRS.3SG seem.PTCP possible to not NEG oppose.INF  
cette conséquence aux entreprises.  
this consequence to DEF.ART companies

‘However, [...] it seemed possible not to oppose this consequence to companies.’

In copulative constructions with an impersonal subject, a dative complement is equally commonly present (156), but not systematically (157).

**(Lack of) agentivity** In both constructions, non-copulative ‘show up’ and copulative ‘seem’, the subject consistently assumes a non-agentive role. The examples hitherto provided have non-agentive subjects, as they lack control or volition over the action, regardless of the construction employed.

This characterization is exemplified by considering instances (158) and (159):

(158) La Palestine géographique m'a parue très belle.  
DEF.ART P. geographical to.me.have.PRS.3SG seem.PTCP.F.SG very beautiful

‘Geographical Palestine seemed very beautiful to me.’

(159) Plus de vingt-cinq titres sont parus et bien d'autres  
over twenty-five titles be.PRS.3PL publish.PTCP.M.PL and well of.others

sont en préparation.  
be.PRS.3PL in pipeline

‘Over twenty-five titles have been published, with many more in the pipeline.’

Both (158) featuring the copulative construction and (159) employing the non-copulative construction portray subjects devoid of agency. The sole disparity lies in the choice of construction.

**Adverbials** Adverbials of temporal and spatial location are the most frequent ones. Interestingly, adverbials of spatial locations (e.g., *en Dakota du Nord* ‘in North Dakota’) are exclusively associated with non-copulative constructions, as illustrated in example (160).

(160) Cet ouvrage est paru [...] dans la collection  
This work be.PRS.3SG appear.PTCP in DEF.ART collection

“Science, histoire et société”.  
 S., H. et S.

‘This work appeared [...] in the “Science, histoire et société” collection.’

The eight instances presenting the spatial location adverbial are exclusively paired with ‘be’ because of their association with the non-copulative construction.

Conversely, the time adverbial can occur with both copulative and non-copulative constructions, but it is notably more frequent in non-copulative constructions, with a total of 25 occurrences, as illustrated in (161).

(161) Trois volumes sont parus en 2008, 2009 et 2017  
three volumes be.PRS.3PL appear.PTCP.M.PL in 2008, 2009 and 2017

aux éditions Dupuis.  
at.DEF.ART publications D.

‘Three volumes appeared in 2008, 2009 and 2017, with Editions Dupuis.’

The example (161) confirms the compatibility of both temporal and spatial location adverbials with the non-copulative construction. The higher prevalence of these adverbials with ‘be’ can be accounted for by the inherent nature of a verb used in its primary meaning ‘appear’ and the closely related and often occurring meaning ‘be published’, which naturally go along with information on date/time and place.

It can be inferred that auxiliary selection for the verb *paraître* is highly conditioned by the type of construction: non-copulative ‘appear/look’ and copulative ‘seem’. While ‘be’ is possible, ‘have’ serves as the primary auxiliary when the verb adopts a copulative construction.

Conversely, ‘be’ takes precedence as the most frequent auxiliary when the verb assumes its prototypical non-copulative construction denoting ‘appearance’. ‘Have’ occurs only occasionally in this context.

Table 4.40 provides a summary of the factors influencing the selection between ‘have’ and ‘be’.

Construction	Human trait	Agentivity	Auxiliary
Non-copulative	Non-human	Not-agentive →	Be
Copulative	Non-human/human	Not-agentive →	Have

TABLE 4.40 – primary factors influencing the selection of ‘have’ and ‘be’ for *paraître*.

Table 4.40 delineates the principal factors shaping the choice of auxiliary verbs. Notably, two distinct constructions correspond to the two different auxiliaries employed.

The non-copulative construction, characterized by non-human and non-agentive subjects, exclusively selects ‘be’, with few exceptions.

In contrast, the copulative construction may feature both non-human and human subjects, non-agentive in nature, and typically opts for ‘have’ as the auxiliary.

#### 4.3.2 *Apparaître*

In FrTenTen (17), *apparaître* ‘appear’/‘seem’ shows a highly predominance of ‘be’, with 1% of occurrences with ‘have’ (198 occurrences) compared to the 99% with ‘be’ (14,090 occurrences) when filtering out instances where a noun appears in the first five tokens to the right.

In contrast to its parasyonym *paraître*, which conveys the notion of ‘be/become visible’, *apparaître* carries a richer lexical meaning, viz. ‘show up’. Similar to *paraître*, *apparaître* can

manifest in two constructions: a non-copulative one, where its lexical meaning remains intact, and a copulative one, where it also functions as an evidential marker, highlighting visual perception.

***Non-copulative vs copulative constructions*** In the 100-sentences sample, both ‘have’ and ‘be’ occur in both constructions available for *apparaître*, as illustrated in Figure 4.31.

In contrast, *paraître*, selects ‘have’ in the copulative construction and ‘be’ in the non-copulative.

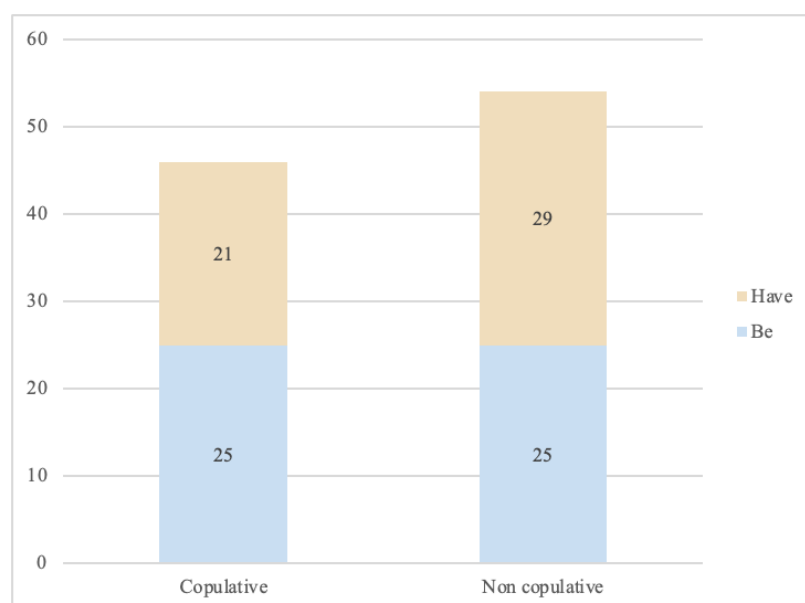


FIGURE 4.31 – Correlation of ‘have’ and ‘be’ with copulative and non-copulative construction.

Factors	‘be’	‘have’	Total
<b>Copulative</b>	25	21	46
<b>Non copulative</b>	25	29	54
<b>Total</b>	50	50	100

TABLE 4.41 – Frequency of ‘have’ and ‘be’ according to the construction.

***Non-copulative constructions*** When the verb is employed in its non-copulative construction, it retains its lexical meaning of ‘show up’ ‘emerge’. Examples (162) and (163) illustrate two instances of the non-copulative construction:

(162) La transplantation [...] est possible dès que les premières  
 DEF.ART transplantation be.PRS.3SG possible as.soon.as DEF.ART first

racines auront apparu.  
 roots have.FUT.3PL appear.PTCP

‘Transplantation [...] is possible as soon as the first roots appear.’

(163) C’est en 2011 que les premiers Chromebooks sont  
 it.be.PRS.3SG in 2011 that DEF.ART first C. be.PRS.3PL

apparus.  
 appear-PTCP.MPL

‘The first Chromebooks appeared in 2011.’

**Copulative construction** The copulative construction, as observed with *paraître*, involves a subject and a predicate and is frequently accompanied by a dative. Referring to table 6.41 just presented, *apparaître* grammaticalized into a copula with the meaning ‘seem’ or ‘appear’, is used with both ‘have’ (164) and ‘be’ (165):

(164) Cette compilation m’ait apparu comme magistrale.  
 This compilation to.me.have.SUBJ.3SG appear.PTCP as masterful

‘[...] this compilation appeared to me as masterful.’

(165) Le rizatripan est apparu plus efficace [...].  
 DEF.ART rizatriptan be.PRS.3SG appear.PTCP more effective

‘Rizatriptan appeared more effective. [...]’

The subjects in these copular constructions are mainly non-human, but there are instances containing human subjects as well.

**Human and non-human subjects** Though human subjects are also present, non-human subjects are much more frequent, as shown in Figure 4.33 and Table 4.42, and show a balanced distribution of ‘have’ and ‘be’ in both constructions.

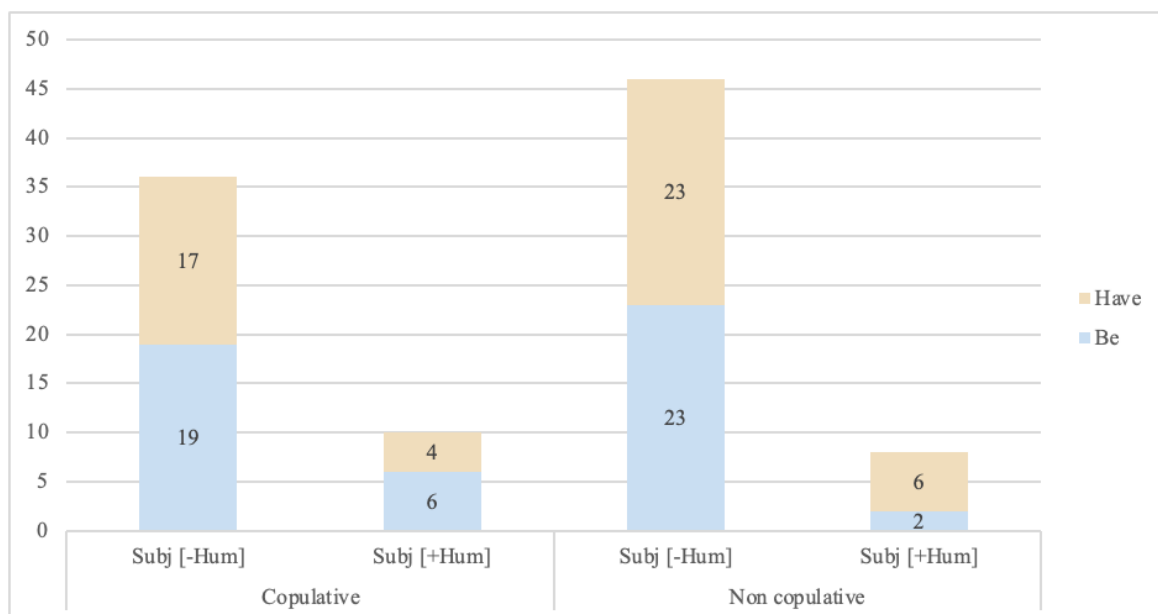


FIGURE 4.32 – Correlation of ‘have’ and ‘be’, the construction’s type and human/non-human nature of the subject.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Copulative Subj [-Hum]</b>	19	17	36
<b>Copulative Subj [+Hum]</b>	6	4	10
<b>Non copulative Subj [-Hum]</b>	23	23	66
<b>Non copulative Subj [+Hum]</b>	2	6	8
<b>Total</b>	50	50	100

TABLE 4.42 – Frequency of ‘have’ and ‘be’ based on the construction’s type and human/non-human nature of the subject.

Surprisingly, the empirical findings reveal that the [ $\pm$ HUMAN] trait of the subject does not exert a significant influence on the choice of auxiliary. Particularly with non-human subjects, there is a balanced distribution: 19 tokens exhibit ‘be’, and 17 exhibit ‘have’ in the copulative construction, while 23 exhibit ‘be’, and 23 exhibit ‘have’ in the non-copulative construction.

These results challenge the assertion made by Legendre and Sorace (2003: 7-8), who suggested that “*apparaître* typically selects *être*, but it is not uncommon to see it with *avoir*”.

The occurrences examined in this corpus, albeit limited to 100 sentences, not only indicate that instances of ‘have’ with *apparaître* are not infrequent but also that they are equally frequent as those with ‘be’.

Human subjects do occur, though less frequently. When the subject is human, ‘have’ is dominant in the non-copulative construction, with 6 occurrences compared to 2 with ‘be’, though the small number of instances prevents definitive conclusions.

Examples for the non-copulative use of ‘have’ and ‘be’ with human subjects are given in (166) and (167), respectively:

(166) Qu’est devenue Sainte Bernadette après que la  
 what.be.PRS.3SG become.PTCP.F.SG S.B. after.that DEF.ART

Vierge Marie lui ait apparu à Lourdes?  
 Virgin Mary to.her have.SUBJ.3SG appear.PTCP in L.

‘What did Saint Bernadette become after the Virgin Mary appeared to her in Lourdes?’

(167) [...] les premiers hominidés sont apparus bien plus tard.  
 DEF.ART first hominids be.PRS.3PL appear.PTCP.M.PL much more late

‘[...] The first hominids appeared much later.’

Conversely, instances with human subjects in the copulative constructions predominantly involve more often ‘be’ (169) than ‘have’ (168). Again, the small number of instances (6 instance of ‘be’ vs 4 instances of ‘have’) prevents to make firm conclusions.

(168) Tout d’abord les guides [...] m’ont apparu comme  
 First.of.all DEF.ART guides [...] to.me.have.PRS.3PL appear.PTCP as

très directifs.  
 very controlling

‘First of all, the guides appeared to me very controlling’.

(169) Au moment de ses auditions, Keller est apparu  
 at.DEF.ART time of his auditions K. be.PRS.3SG appear.PTCP

“un peu agité.”  
 a.little agitated

‘At the time of this auditions. Keller appeared a little agitated’.

Two notable features distinguish *apparaître* from *paraître*. Firstly, both the plain lexical construction and the copulative construction are compatible with human subjects. In the former construction, *apparaître* retains its general sense of ‘be manifest’ or ‘showing up’, without acquiring the specific meaning of ‘be published’.

Secondly, concerning auxiliary selection, there is less apparent alignment between ‘have’ and the copula construction, and between ‘be’ and its use as a full lexical verb.

**(Lack of) agentivity** In line with the analysis conducted for *paraître*, non-human subjects predominantly entail non-agentive subjects in both the copulative and non-copulative constructions (Figure 4.33):

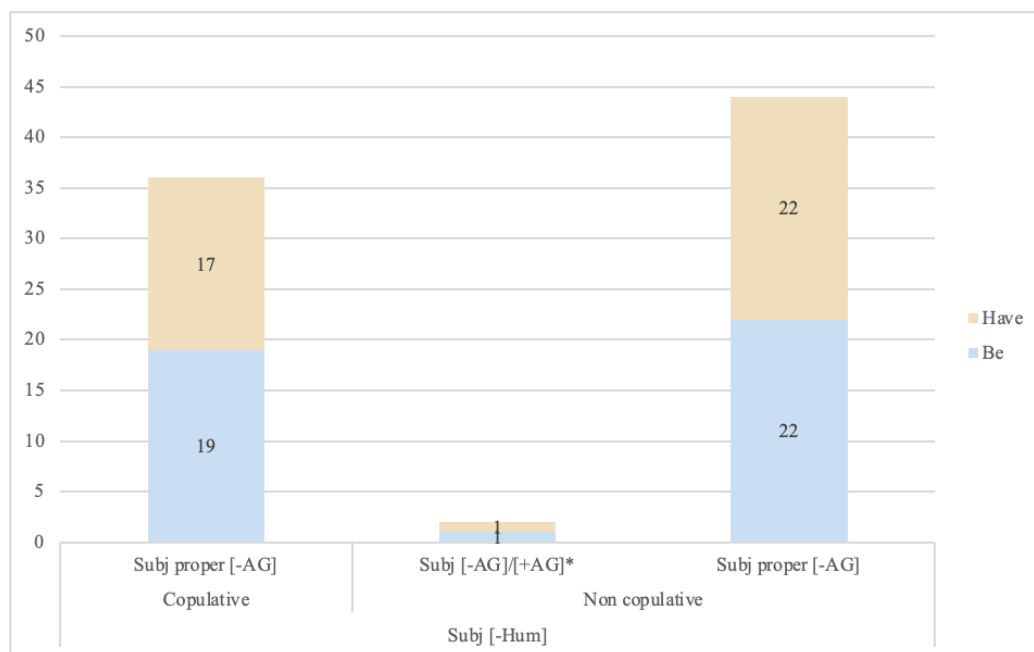


FIGURE 4.33 – Correlation of ‘have’ and ‘be’, the construction’s type and non-human subjects.

Factors	Data ‘Be’	Data ‘Have’	Total
<b>Copulative Subj proper [-AG]</b>	19	17	36
<b>Non Copulative Subj [-AG] / [+AG] *</b>	1	1	2
<b>Non-Copulative Subj proper [-AG]</b>	22	22	44
<b>Total</b>	42	40	82

TABLE 4.43 – Frequency of ‘have’ and ‘be’ based on construction’s type and semantic features of the subject (non-human and internal cause/not agentive subjects).

In consideration of the verb’s semantics, the aspect of internal cause does not exert a substantial influence. Apart from two instances where non-human subjects act as internal causes, all other occurrences involve non-agentive subjects. The presence of internal cause is more prevalent with human subjects, as depicted in Figure 4.34.

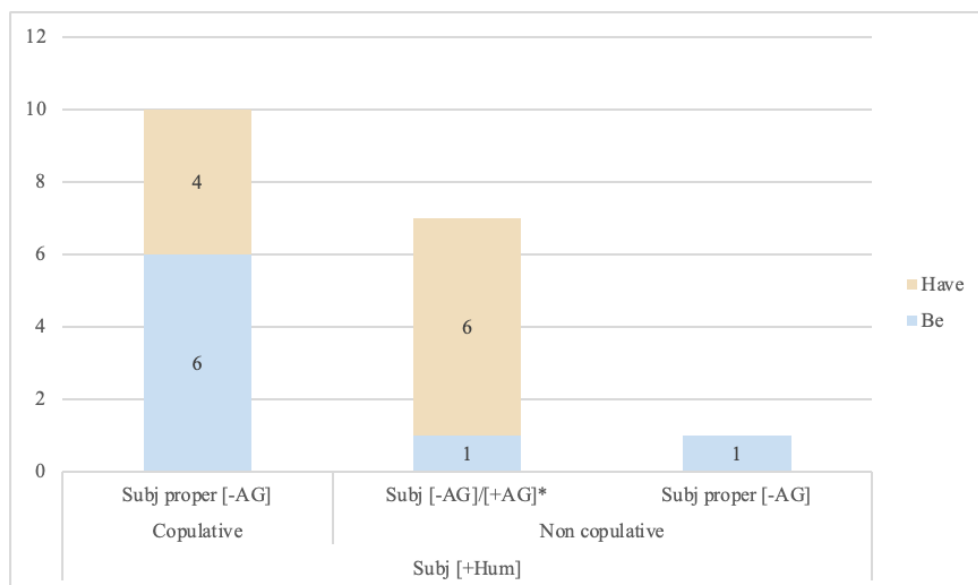


FIGURE 4.34 – Correlation of ‘have’ and ‘be’, the construction’s type and human subjects (internal cause and non-agentive).

Factors	Data ‘Be’	Data ‘Have’	Total
Subj [+Hum] Copulative Subj proper [-AG]	6	4	10
Subj [+Hum] Non Copulative Subj [-AG] / [+AG]*	1	6	7
Subj [+Hum] Non-Copulative Subj proper [-AG]	1	/	1
<b>Total</b>	8	10	18

TABLE 4.44 – Frequency of ‘have’ and ‘be’ based on construction’s type and semantic features of the subject (human and internal cause/not agentive subjects).

Even though the numbers are low, instances of internal cause exclusively occur in the non-copulative construction, preserving the verb’s inherent meaning. Examples are (170) and (171), where the subject acts as an internal cause, as it contributes to the process of emerging, but is devoid of volition.

(170) Quant aux Pénitents de la [...] compagnie des Pénitents Bleus,  
As to.DEF.ART penitents of DEF.ART compagnie of.DEF.ART P.B.

ils ont apparu en 1040 [...].  
they have.PRS.3PL appear.PTCP in 1040

‘As for the penitents of the [...] compagnie of the Pénitents Belus, they appeared in 1040.’

(171) Un jour, alors que j’apprenais à prier, il est apparu.  
INDF.ART day as I.learn.IPFV.1SG to pray he be. PRS.3SG appear.PTCP

‘One day, as I was learning to pray, he appeared.’

**Adverbials** The most frequent adverbials observed express temporal and spatial location. The spatial location adverbial is exclusively linked with the non-copulative construction, with 6 instances featuring ‘have’ and 3 with ‘be’.

Conversely, the time adverbial appears in both constructions, albeit more frequently in the non-copulative one. Despite the constrained number of occurrences and the balanced distribution, it can be inferred that these adverbials do not exert a significant influence on the selection of the auxiliary.

In conclusion, the verb *apparaître* displays a unique behavior, showcasing a balanced distribution between ‘have’ and ‘be’ in both copulative and non-copulative constructions. While the non-copulative construction marginally influences the choice of auxiliary when the subject involves a human and internal cause, this effect appears relatively moderate.

Overall, the variability of auxiliaries with this verb is well-documented, and the grammatical commentary in (172) offers a concise summary of the observed patterns with this particular verb:

(172)	Il	se	conjugue	avec	l’auxiliaire	<i>avoir</i>	et	l’auxiliaire
	3SG	REFL	conjugate.PRS.3SG	with	DEF.ART.auxiliary	have.INF	and	DEF.ART.auxiliary
	<i>être</i> :	ces	spectres	m’ont		apparu	ou	me
	be.INF	these	specters	to.me.have.PRS.3PL		appear.PTCP	or	to.me
	sont		apparus [...]					
	be.PRS.3PL		appear.PTCP.M.PL					

‘It is conjugated with the auxiliary to have and the auxiliary to be: these specters have appeared to me or have appeared to me [...].’

The sentence in (172) aptly illustrates the uncertainties inherent in verbs allowing auxiliary alternation, as both ‘have’ and ‘be’ are employed. This sentence originates from FrTenTen (17), with a dictionary serving as a point of reference.<sup>25</sup>

In Table 4.45, the primary factors co-occurring in the selection of the auxiliary are summarized.

---

<sup>25</sup> The dictionary is *sensagent*: [dictionnaire.sensagent.leparisien.fr](http://dictionnaire.sensagent.leparisien.fr).

Construction	Human trait	Agentivity	Auxiliary
Non-copulative	Human	Internal cause →	Have
Non-copulative	Non-human	Non-agentive →	Have/Be
Copulative	Human/Non-human	Non-agentive →	Have/Be

TABLE 4.45 – Primary factors co-occurring in the choice of ‘have’ and ‘be’.

The analysis of the verb *apparaître* presented in table 4.45 reveals three predominant scenarios. Firstly, the verb demonstrates two distinct constructions: copulative and non-copulative.

In copulative constructions, the subject can be either human or non-human, consistently lacking agentivity, and the distribution of ‘have’ and ‘be’ is balanced (with a slight preference for ‘be’).

Secondly, in non-copulative constructions with a non-human subject, the choice between ‘have’ and ‘be’ is evenly distributed.

The third scenario, less frequent, involves a non-copulative construction with a human subject, often serving as an internal cause. The auxiliary ‘have’ is more frequent in this context, especially because of the internal cause factor.

#### 4.3.3 Conclusion: Verbs of Appearance

The third group of verbs analyzed in this section 4.3 comprises the verbs of appearance, where the construction plays the most influential role: when the verb is utilized in its non-copulative construction, the auxiliary mostly used is ‘be’: *paraître* prefers ‘be’, while *apparaître* presents both ‘have’ and ‘be’.

Conversely, in the copulative construction, the auxiliaries used are ‘have’ for *paraître* and ‘have’ and ‘be’ for *apparaître*. It can be deduced that the auxiliary selection with *apparaître* is not influenced by the construction (Table 4.46).

Most influential factors for <i>Verbs of Appearance</i>	
CONSTRUCTION	AUXILIARY
Non-copulative	Have/Be
Copulative	Have/Be

TABLE 6.46 - Most influential factors for *Verbs of Appearance*.

In the non-copulative construction, the auxiliary ‘have’ is used when the subject is human and an internal cause, whereas a preference for ‘be’ is observed when the subject is non-agentive. However, the data involving human subjects are insufficient to establish a clear tendency.

#### 4.4 CONCLUDING REMARKS ON AUXILIARY SELECTION IN FRENCH

The examination of auxiliary variation in French peripheral verbs reveals a non-uniform picture. This complexity arises from the fact that the verbs under consideration fall into three main types, viz. spatial motion verbs, change-of-state verbs, and verbs of appearance, and that each type exhibits distinct patterns of auxiliary usage, without the possibility of establishing a common principle that governs the selection of auxiliaries. A particular case is *diminuer*, a change-of-state verb that behaves similarly to the spatial motion verb *baisser*.

The spatial motion verbs *passer*, *monter*, *baisser*, and *descendre* display different constructions, including spatial motion and quantitative evolution. In constructions with a primary spatial motion meaning, ‘be’ is generally preferred, resembling core verbs like *partir* ‘leave’ or *venir* ‘come’. When these verbs appear in a quantitative evolution construction, the conveyed meaning is derived, and ‘have’ tends to be favored, particularly in instances with non-human subjects. The change-of-state verb *diminuer* aligns with the quantitative evolution construction.

The change-of-state verbs *mûrir*, *diminuer*, *changer*, and *exploser* may involve the internal cause factor and the lack of agentivity influencing auxiliary selection. The internal cause factor is especially important for *mûrir* and *changer*. If the subject is an internal cause, the preferred auxiliary is ‘have’. *Mûrir* predominantly features internal cause subjects due to its semantics. Otherwise, with non-agentive subjects, the favored auxiliary is ‘be’ (especially for *changer* and *mûrir*). *Exploser* represents a particular case of change-of-state verbs as it presents different constructions. When the verb is used in its literal core meaning delineating a process, both ‘have’ and ‘be’ can be equally used. When it is embedded in a quantitative evolution construction, the auxiliary is ‘have’; when it is in a semi-auxiliary or lexicalized construction, the auxiliary is ‘be’. In all instances, the subjects are non-agentive.

The verbs of appearance, *paraître* and *apparaître*, exhibit two major constructions, copulative and non-copulative constructions. *Paraître* exhibits a clear and distinct pattern: ‘have’ with the copulative construction and ‘be’ with the non-copulative. In contrast, *apparaître* presents both auxiliaries with both constructions, demonstrating less influence by construction compared to *paraître*.

Another relevant distinction for auxiliary selection for spatial motion and change-of-state verbs is the *process vs. resultative state* dichotomy. Verbs compatible with both ‘have’ and ‘be’ auxiliaries may choose ‘be’ to highlight the resulting state. However, this pattern is ambiguous for labile verbs, where the ‘be’ + past participle pattern may be analyzed either as an intransitive construction or as passive voice construction of the corresponding transitive construction. The high frequency of the ‘have’ for intransitive constructions in French, associated with a processive reading, often explains why verbs allowing both auxiliaries are frequently interpreted as resultative if they select ‘be’.

Table 4.47 summarizes the key factors influencing the distribution of auxiliaries in French peripheral verbs. These factors encompass the nature of the verb, the presence of spatial motion constructions, the role of internal cause subjects, and the consideration of processive vs. resultative aspect. The complex interplay of these factors contributes to the varied usage of ‘have’ and ‘be’ auxiliaries across different verb types and constructions.

General Factors	Type of Factor	Different types of verbs
<b>1: CONSTRUCTION</b>	<i>Spatial vs. Quantitative construction</i>	<i>Spatial</i> : ‘be’ <i>Quantitative</i> : ‘have’
	<i>Non copulative vs. Copulative construction</i>	<i>Copulative</i> : ‘have’ and sometimes ‘be’ <i>Non copulative</i> : ‘be’ and sometimes ‘have’ based on other factors
<b>2: HUMAN TRAIT</b>	<i>Animacy and human trait</i>	[+Hum] for <i>Spatial motion</i> verbs: ‘be’ [+Hum] for <i>Change-of-state</i> verbs: ‘have’ [+Hum] for <i>Verbs of appearance</i> : ‘have’ for <i>paraître</i> , ‘have/be’ for <i>apparaître</i>
<b>3: AGENTIVITY</b>	<i>Internal causative behavior</i>	[+AG], [+Internal Cause] for <i>Spatial motion</i> verbs: ‘be’ [+AG], [+Internal Cause] for <i>Change-of-state</i> verbs: ‘have’
	<i>Lack of agentivity</i>	[-AG] for <i>Spatial motion verbs</i> : ‘have’ (quantitative evolution construction) [-AG] for <i>Change-of-state verbs</i> : ‘be’ [-AG] for <i>Verbs of appearance</i> : ‘have’/‘be’
<b>4: ASPECT</b>	<i>Resultative interpretation</i>	Result: ‘be’ Process: ‘have’
<b>5: VOICE</b>	<i>Passive construction</i>	Active: ‘have’ Passive: ‘be’

TABLE 4.47 – List of aspects that could affect the distribution of the auxiliaries.

The selection of auxiliaries in French verbs is a complex interplay of various factors. A simple list of the factors at work, no matter how complete, is not sufficient to explain the selection of auxiliaries. It is important to prioritize them and measure their respective impact, which will be the aim of the next chapter.

For some verbs, primary determinants encompass the specific construction in which the verb is employed, whether it pertains to spatial motion construction, a quantitative evolution construction or has a temporal meaning. Within the realm of construction factors, distinctions arise between copulative and non-copulative constructions (as observed in certain verbs, such as verbs of appearance). Interestingly, a certain type of construction, e.g. the quantitative evolution construction of the spatial motion construction, may be associated to the same auxiliary selection principles for several verbs. However, as has been shown for *paraître* and *apparaître*, this is not always the case.

Factors such as the human trait or the presence of internal cause subjects may influence auxiliary selection for specific verbs. For some change-of-state verbs, the [+HUMAN] and [+INTERNAL CAUSE] or [+AGENTIVE] factors are influential, while for spatial motion verbs, these semantic attributes are only derived traits. For example, if the construction involves spatial motion, the human subject will select ‘be’, not due to the human trait, but because of the construction.

Finally, aspect and voice play a role in auxiliary selection, insofar as ‘be’ is associated to resultative aspect and to passive voice. It is important to remember that the pattern *être* + *participe passé* is sometimes liable to multiple analyses (intransitive, passive, adjectival), without it always being possible to decide definitively (Buchard & Carlier 2008)



## CHAPTER 5: STATISTICAL ANALYSIS

This chapter presents the outcomes of statistical analyses based on two distinct methodologies: CHAID (Chi-squared Automatic Interaction Detector) (Kass 1980) and Random Forest (Breiman 2001; Liaw & Wiener 2002).

The utilization of these statistical methods represents a novel approach in auxiliary selection analysis. To the best of my knowledge, these methods have not yet been used to validate and interpret hypotheses on auxiliary selection derived from corpus analysis.

CHAID and Random Forest have played a pivotal role in substantiating my hypotheses and elucidating relationships within the dataset. The convergence of results obtained from both statistical methods serves as a robust validation of the proposed hypotheses. These two statistical methods have yielded noteworthy insights, particularly concerning Italian verbs exhibiting auxiliary alternation. Notably, the analysis corroborated that internal cause (Mateu 2009) serves as the primary influencing factor for Italian verbs, while the distinction among French verbs is contingent upon verb types and verb constructions.

Subsequent to the introduction of CHAID and Random Forest methods in Sections 5.1 and 5.2, Sections 5.3, 5.4, and 5.5 will delve into the analysis and outcomes pertaining to the methodologies applied to the verbs analyzed in this study: Italian intransitive verbs having the status of main verbs, Italian aspectual semi-auxiliary verbs, and French intransitive verbs, respectively. Within Section 5.3, the different classes of French verbs will be analyzed separately, with 5.3.1 focusing on spatial motion verbs, 5.3.2 on change-of-state verbs, and 5.3.3 on verbs of appearance. Finally, Section 5.6 will offer concluding remarks.

### 5.1 INTRODUCTION: CHAID

CHAID is a decision tree algorithm used for predictive modeling and classification. It employs a chi-square splitting criterion. The key features of CHAID are outlined by Ritschard (2013: 9-10):

- 1. “At each node, CHAID determines, for each potential predictor, the optimal  $n$ -ary split it would produce, selecting the predictor based on these optimal splits.”
- 2. “CHAID utilizes  $p$ -values with Bonferroni correction as splitting criteria.”

Utilizing  $p$ -values as criteria for growth integrates stopping rules that inherently account for statistical significance. These rules set thresholds at customary critical values, such as 1%, 5%, or 10%, responsive to the sample size in each split, thus preventing segmentation into excessively small groups (see Ritschard 2013 for more details).

In the context of this study, CHAID aims to construct a tree structure predicting the target variable (e.g., the selected auxiliary) on the basis of the predictor variables' values (e.g., the semantic and syntactic parameters used in the corpus analysis).

The algorithm iteratively forms subsequent smaller sub-groups by maximizing differences between segments and minimizing differences within a given segment (e.g., it groups together similar sentences according to the parameters), selecting the strongest predictor capable of splitting a node.

Kass (1980:2) defines the method as follows:

“CHAID proceeds in steps. First, the best partition for each predictor is found. Then, the predictors are compared, and the best one is chosen. The data are subdivided according to this chosen predictor. Each of these subgroups is re-analyzed independently to produce further subdivisions for analysis. The type of each predictor determines the permissible groupings of its categories, optimizing the contingency table's significance level according to the chi-squared test. [...] This ensures sufficient observations to validate the test”.

What sets CHAID apart from previous statistical tools is Bonferroni's correction to  $p$ -values (Ritschard 2013: 20). The Bonferroni adjustment addresses the issue of multiple testing. For instance, if a  $p$ -value for a specific split is deemed statistically significant, all other potential splits'  $p$ -values should exceed this threshold. To accommodate multiple comparisons, the Bonferroni correction reduces the critical value for each test by dividing it by the total number of tests or, in CHAID's case, by multiplying the  $p$ -value of the optimal solution by the number of tests.

I selected this classification model because it offers criteria to create homogeneous groups of sentences by maximizing inter-variability between groups and minimizing intra-variability within groups. This method provides a solid foundation for demonstrating the empirical validity of my hypotheses in a grounded manner.

## 5.2 INTRODUCTION: RANDOM FOREST

Random Forest is an ensemble learning method based on decision trees, first proposed by Breiman (2001). As stated by Liaw & Wiener (2002: 18), “in standard trees, each node is split using the best split among all variables. In a random forest, each node is split using the best among a sub-set of predictors randomly chosen at that node.”

The algorithm functions as follows: before constructing each decision tree, the bootstrapping process randomly selects a subset of the training data. At each node of the decision tree, only a random subset of features is considered for splitting, and the best split is chosen from these random variables (see Liaw and Wiener 2002: 18). Subsequently, each decision tree in the Random Forest ensemble generates an output, and the random forest “predicts new data by aggregating the predictions of the trees” (Liaw and Wiener 2002: 18).

Randomness is a fundamental aspect of Random Forest, as it enhances the model’s generalization performance by reducing the variance of individual trees. This variance reduction is achieved by training multiple trees on different subsets of data and features.

Breiman (2001: 29) underscores this effectiveness:

“Random forests are an effective tool in prediction. Because of the Law of Large Numbers, they do not overfit. Injecting the right kind of randomness makes them accurate classifiers and regressors. Furthermore, the framework in terms of strength of the individual predictors and their correlations gives insight into the ability of the random forest to predict.”

In other words, RF gives us a graph complimentary to the decision tree given by CHAID. RF provides a list of the parameters in a decreasing order, allowing to appreciate the weight of each parameter. These two graphs show us two different pictures of the same phenomenon: CHAID groups similar sentences according to the best predictor and the subsequent one that is constrained by the previous, while RF gives us a score for every predictor.

## 5.3 CHAID AND RANDOM FOREST FOR ITALIAN VERBS

**CHAID** The CHAID analysis was performed using SPSS version 26 and STATA software version 15. As pointed out in Section 5.1, CHAID was employed to predict the auxiliary (target variable) based on the values of predictor variables, comprising sixteen parameters. These parameters were identified during the corpus analysis. Four of them relate to the semantic nature of the subject (human trait, animacy, agentivity, and internal cause) and twelve other parameters

concern the presence of adverbials and their nature (argument adverbial, adverbial of manner, quantity adverbial, adverbial indicating telic aspect, atelic aspect, or progressive aspect, adverbial indicating temporal location, or temporal location combined with duration, adverbials indicating spatial location, direction with endpoint or direction without endpoint, and absence of adverbial).

For the Italian intransitive verbs, allowing both auxiliaries, the algorithm initially identified internal cause as the primary splitting factor, subsequently recognizing human trait and quantity adverbial as significant contributors. Figure 5.1 depicts the decision tree, which visualizes the hierarchy of parameters conditioning the choice of the auxiliary for main verbs in Italian.

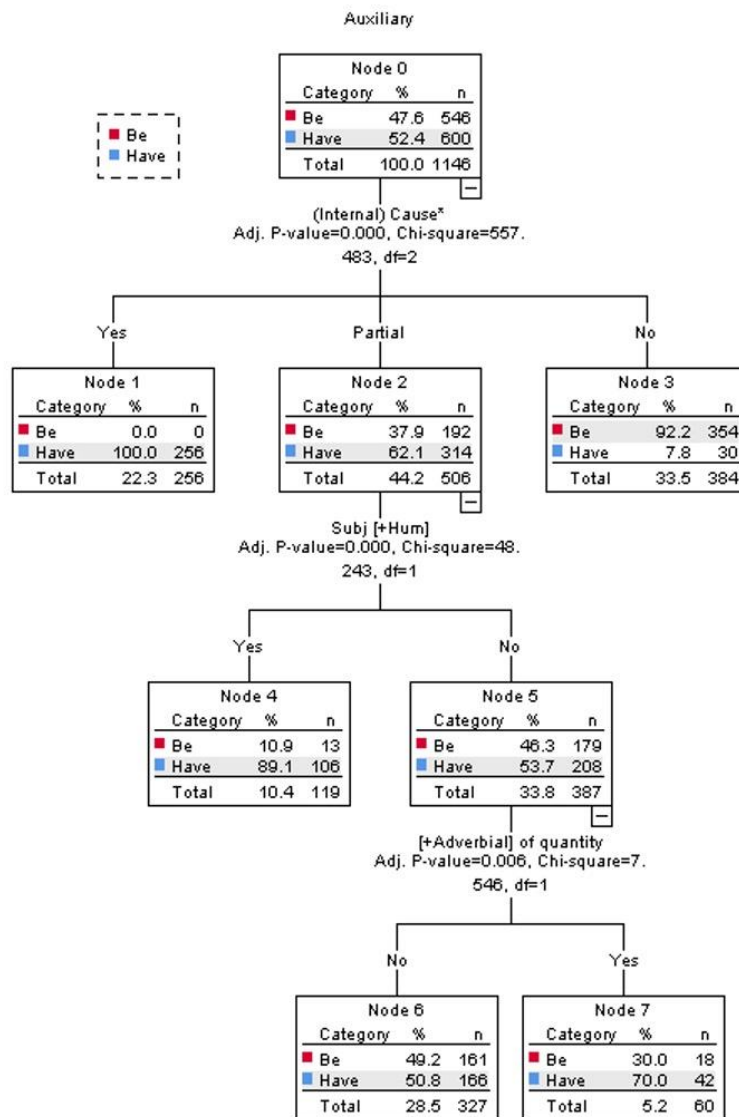


FIGURE 5.1 – Decision tree of the 13 Italian verbs.

The initial node (Node 0) in the analysis represents the entire corpus of intransitive verbs, comprising 1146 sentences containing 13 different lexemes. It undergoes division into three sub-nodes based on what CHAID identifies as the most significant predictor ( $p$ -value  $< 0.0001$ ) of auxiliary selection for Italian main intransitive verbs, namely internal cause.

The parameter internal cause was conceived in terms of three variables because it was otherwise impossible for the software to recognize that the internal cause subject could be both the agent and the patient of the action. Therefore, during the CHAID analysis, a value of 0 was assigned to the absence of agentivity, 1 to the presence of internal cause (equivalent to ‘partial’ in the figure 5.1, therefore Node 2), and 2 for full agentivity.

Within Node 1, where the subject exhibits clear agentivity, all sentences feature the auxiliary ‘have’ (100%). Conversely, in instances where internal cause is absent and the subject lacks agentivity, ‘be’ is selected in approximately 92.2% of sentences (Node 3).

Node 2 is characterized by internal cause without agentivity. This combination of features leads to a distribution of auxiliaries of approximately 60/40%, with a higher frequency of ‘have’. Notably, Node 2 contains a significantly higher number of occurrences than Nodes 1 and 3.

Thanks to this heterogenous situation, CHAID can identify another highly significant parameter (Subj[+Hum]) with a  $p$ -value  $< 0.0001$ , which further splits Node 2. In this division, subjects with the [+Hum] attribute predominantly select ‘have’ (89.1%) (Node 4), while those with [-Hum] attributes exhibit a nearly equal distribution between the two auxiliaries (Node 5).

CHAID then further splits Node 5 instead of Node 4 because of two reasons: Node 5 contains three times more sentences than Node 4 (387 vs 106), and Node 5 is more heterogenous than Node 4 (where almost 90% of sentences contains ‘be’). It retrieves another attribute capable of drawing a division between this group of sentences and found the presence/absence of adverbial of quantity as the best predictor.

It thus forms Node 6 (absence of quantity adverbials) and Node 7 (presence of quantity adverbials). Despite the size and heterogeneity of Node 6, the sentences in this branch are not characterized by statistically significant differences between them. In other words, none of the 16 attributes grounds any possible further grouping of sentences (a point that can be partially confirmed by the list created by Random Forest, in which all the attributes below the adverbial of quantity are clearly not of interest).

Concerning the feature of adverbials of quantity, the presence of adverbials such as *molto*, *tanto* (Node 7) increases the likelihood of selecting ‘have’ (70%). This tendency may be attributed to the possibility that quantity adverbials are being (mis)interpreted as object

complements, as observed with the verb *cambiare*, yielding the selection of ‘have’ characteristic of transitive verbs.

Three factors therefore prove to be relevant, namely – in decreasing order of importance – internal cause (also implying the presence or absence of agentivity), the presence of the [+HUMAN] attribute, and the presence of quantity adverbials.

CHAID not only identifies the most significant factors based on the significance of *p*-values but also predicts the percentage of correctly classified cases (f-score). The overall percentage of correct predictions is 80%, as depicted in Table 5.1.

Classification			
	Predicted		
Observed	Be	Have	Percent Correct
Be	354	192	64,8%
Have	30	570	95,0%
Overall Percentage	33,5%	66,5%	80,6%

TABLE 5.1 – Classification table for CHAID model (Italian verbs)

Specifically, in 80,6% of cases, three of the parameters assigned to the sentences in the sample enabled CHAID to accurately predict the choice of auxiliary for a given verb in a given sentence from the list. This suggests that the sixteen parameters examined in this study can account for more than eighty percent of the variation in auxiliary selection among the verbs analyzed in this study (as shown in 5.1).

In essence, the process of recursive partitioning conducted by CHAID aligns with my initial intuition, and the quantitative results reinforce the qualitative interpretation regarding auxiliary selection. Moreover, CHAID allows us to establish a hierarchy of the sixteen factors influencing auxiliary selection, as identified through corpus analysis. This statistical method enabled us to determine the most significant factors.

To conclude, the overall percentage of correctly classified cases demonstrates the validity of the model proposed in this study.

**Random Forest** To offer an alternative perspective on the importance of each predictor in my analysis, I applied to the same dataset the Random Forest (RF) model algorithm, developed by Breiman (2001) and Liaw & Wiener (2002). RF assesses the significance of various predictors in determining the target variable (auxiliary), shedding light on which variables exert

the greatest influence on classification. This analysis was conducted using the R-package “caret” version 6.0 (Kuhn 2008). The results, depicted in Figure 5.2, highlight that among the 16 parameters, four have emerged as the most crucial: internal cause, human trait, animacy, and agentivity (or lack thereof).

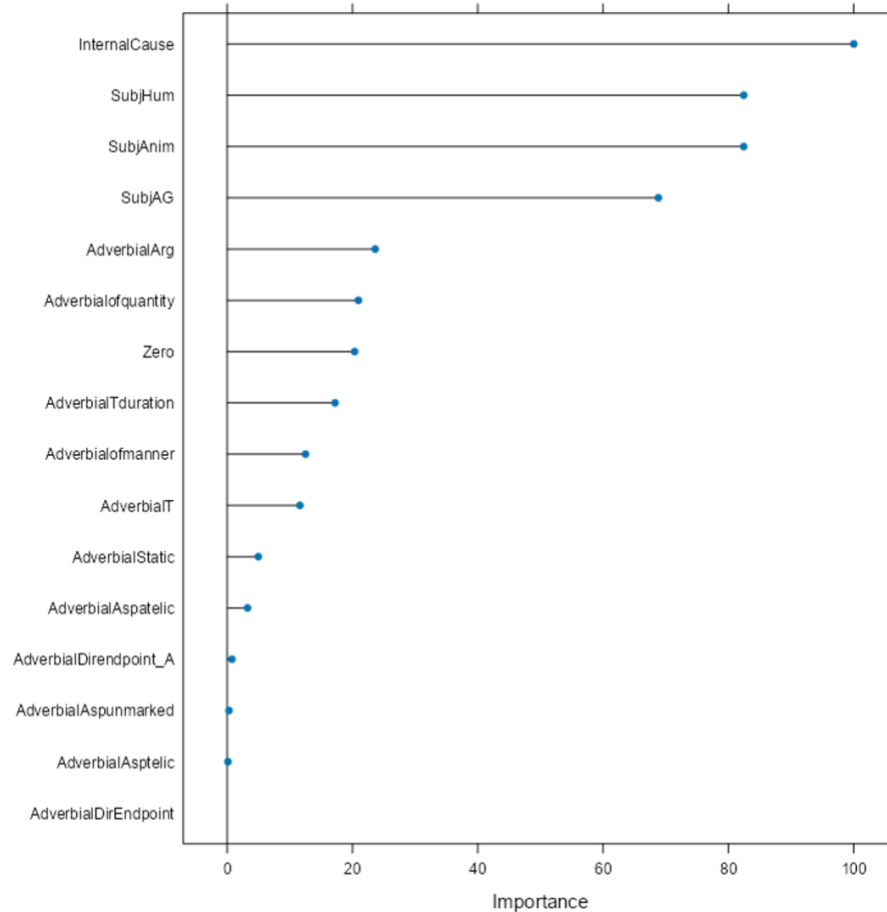


FIGURE 5.2 – Ranking of the sixteen parameters as revealed by RF (Italian verbs)

Agentivity occupies the fourth position in the ranking, although it is worth noting that the internal cause factor already encompasses this parameter, as it constitutes a three-variable parameter.

In my analysis, both agentivity and internal causation have been taken into account, dividing these into three categories: fully agentive subjects, fully non-agentive subjects, and subjects that can simultaneously be agents and patients (internal cause). This distinction was crucial for accurately identifying ambiguous cases and analyzing them with software that detects absence of internal cause (= ‘no internal cause, no agentivity), partial presence (= ‘internal cause, no agentivity), and full presence (= internal cause, agentivity).

It is notable that semantic parameters defining the nature of the subject hold higher positions in the ranking compared to the presence or absence of adverbials and their nature. This predominance of semantic parameters may be attributed to the unique nature of these verbs. This does not necessarily imply that semantics universally play a more significant role, but rather reflects the specific characteristics of peripheral verbs analyzed in this study. The distribution of auxiliaries in these cases likely hinges on subtle semantic nuances compared to core verbs, where syntactic contexts such as endpoints (conveying telicity) may exert greater influence.

In conclusion, both CHAID and Random Forest (RF) are complementary methods and equally important. Considering Random Forest is crucial because CHAID may overlook significant factors due to its segmentation process. CHAID effectively identifies the most significant factors and the percentage of occurrences of ‘have’ and ‘be’ based on these factors. In contrast, RF provides a hierarchy of all factors, assigning a weight of influence to each.

Comparing CHAID and Random Forest offers a comprehensive view of the most important factors in auxiliary selection. Both methods agree that the primary factors influencing auxiliary selection for Italian simple verbs are internal cause, (lack of) agentivity, and the human trait.

Table 5.2 summarizes the main significant factors identified in both CHAID and RF.

Most significant factors affecting auxiliary selection		
	<b>CHAID</b>	<b>RF</b>
1	<b>Internal Cause</b>	<b>Internal Cause</b>
2	<b>[±HUMAN]</b>	<b>[±HUMAN]</b>
3	<b>Quantity Adverbial</b>	Animacy
4		Agentivity
5		Argument adverbial
6		<b>Quantity Adverbial</b>

TABLE 5.2 – Most significant factors influencing auxiliary selection as identified by CHAID and RF (Italian verbs)

Internal Cause and Human Trait emerged as the first two significant factors for both CHAID and RF. The third significant factor is the quantity adverbial for CHAID.

The distinction between CHAID and RF lies in their methodologies: CHAID operates by splitting nodes based on *p*-values, thereby identifying the most important nodes based on previously split variables.

In contrast, RF provides a measure of feature importance, reflecting the relative contribution of each predictor to the overall predictive accuracy of the model. While Random Forest does

consider random subsets of predictors at each split, it does not output a ranking of all factors in the same way CHAID does.

Instead, it focuses on determining the best splits for each decision tree individually, and the importance score reflects the cumulative effect of all splits involving a particular predictor.

Consequently, Random Forest may not always produce the same ranking of factors as CHAID, as it evaluates predictors based on their predictive power within the context of the entire ensemble of decision trees. The first adverbials to appear are the argument adverbials and the quantity adverbial, which are also deemed relevant by the CHAID method.

#### 5.4 CHAID AND RANDOM FOREST FOR ITALIAN ASPECTUAL VERBS

**CHAID** CHAID performance with Italian aspectual (or semi-auxiliary) verbs shows that the most significant factor, with a  $p$ -value  $< 0.001$ , is the type of infinitive, as shown in Figure 5.3. This result corroborates the analysis conducted in Chapter 3.

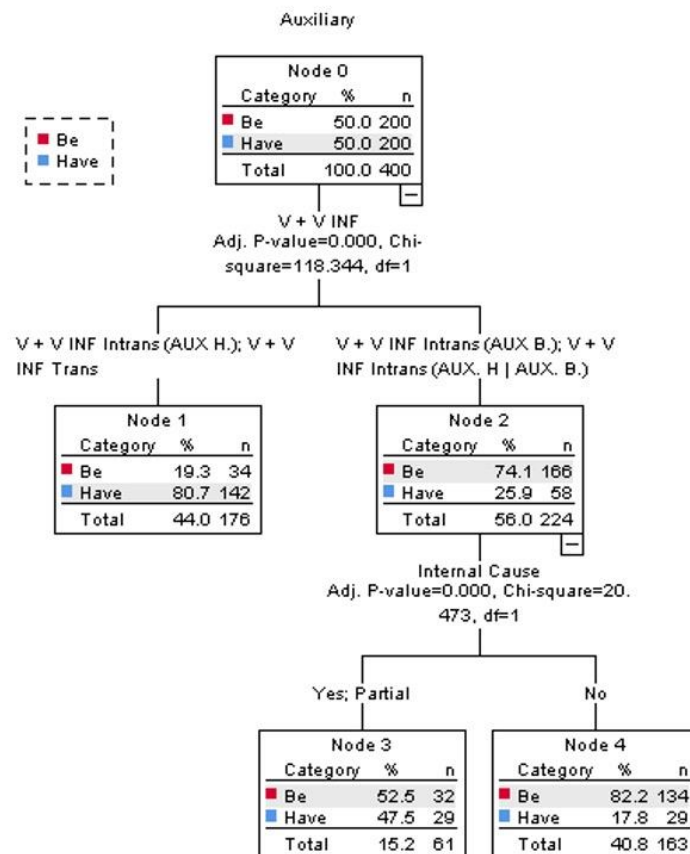


FIGURE 5.3 – Decision tree of the 4 Italian semi-auxiliary verbs.

Node 1 shows the percentage of ‘have’ and ‘be’ when the infinitive is a transitive or an intransitive ‘have’ selecting verb: in this case, the auxiliary ‘have’ is selected in 80.7% of occurrences compared to 19.3% for ‘be’.

Node 2 shows the presence of an intransitive infinitive, either ‘be’ selecting or selecting both auxiliaries: in this case, ‘be’ is the favored auxiliary, with a 74.1% frequency compared to 25.9% for ‘have’. However, when the infinitive is represented by an intransitive ‘be’ selecting verb or a verb selecting both ‘have’ and ‘be’, not only is the infinitive important, but also the type of subject can lead to another split: the subject as an internal cause, agentive, or non-agentive.

Node 3 presents instances with agentive or internal cause subjects, as CHAID has not detected a significant distinction between the two. In this case, even if the infinitive is a ‘be’ selecting verb (or selects both auxiliaries), due to the agentivity or internal cause of the subject, the percentage of ‘have’ increases. There is a 52.5% frequency of ‘be’ compared to 47.5% for ‘have’.

In contrast, when the subject is non-agentive (Node 4) and the infinitive is an intransitive be-selecting verb or both, the favored auxiliary continues to be ‘be’, with a frequency of 82.2% for ‘be’ versus 17.8% for ‘have’.

For Italian semi-auxiliary verbs of aspect, the overall percentage of correctly classified cases is high: the parameters assigned to the sentences in the sample enabled CHAID to accurately predict the choice of auxiliary in 77.0% of the cases, as shown in Table 5.3.

<b>Classification</b>			
	Predicted		
Observed	Be	Have	Percent Correct
Be	166	34	83,0%
Have	58	142	71,0%
Overall Percentage	56,0%	44,0%	77,0%

TABLE 5.3 – Classification table for CHAID model (Italian semi-auxiliary verbs)

The data in Table 5.3 indicates that in 77% of cases, the parameters used in my study—namely, the type of infinitive and the semantic features—can predict the choice of auxiliary.

**RANDOM FOREST** The analysis conducted through the complementary method of Random Forest (RF) confirms that the type of infinitive, followed by the semantic features of the subject, are significant, as shown in Figure 5.4.

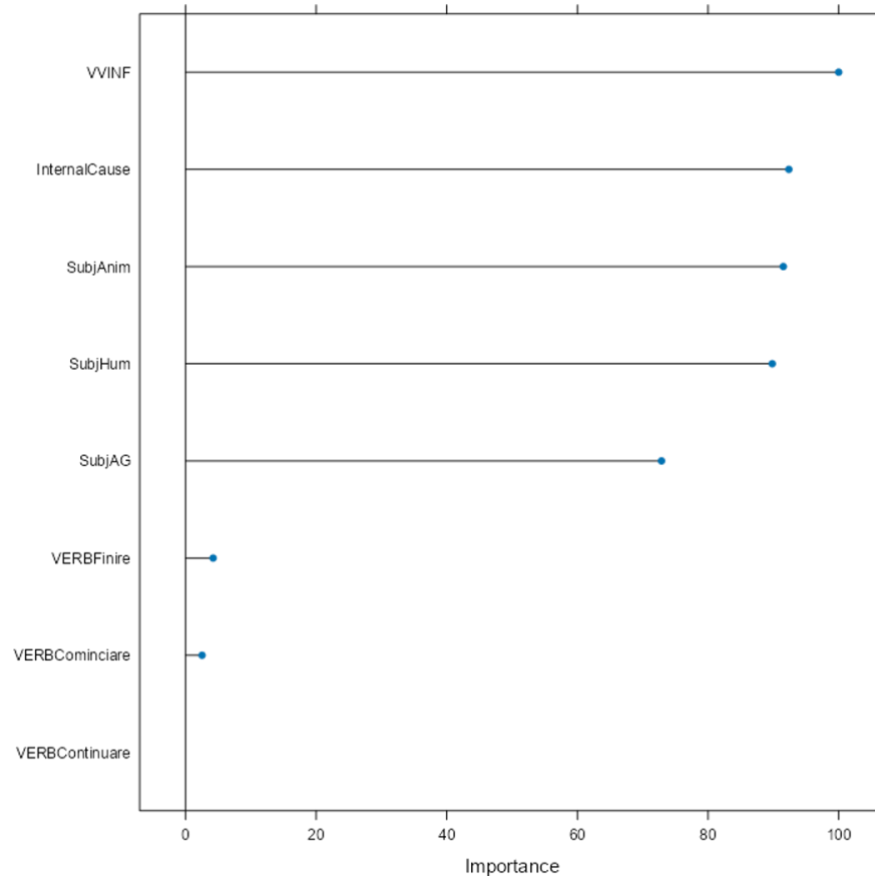


FIGURE 5.4 – Ranking of the sixteen parameters as revealed by RF (Italian semi-auxiliary verbs).

In the hierarchy provided by the Random Forest (RF), the type of infinitive is in the first position, immediately followed by internal cause (including agentivity and lack of agentivity, as it is a three-variable parameter in this analysis) and semantic features such as animacy and human trait.

It is possible to conclude that the combination of the type of infinitive and semantic parameters provides a comprehensive explanation of auxiliary selection in semi-auxiliary verbs.

Table 5.4 provides a summary of the most significant factors identified by both algorithms, CHAID and RF.

Most significant factors affecting auxiliary selection		
	CHAID	RF
1	<b>Infinitive Type</b>	<b>Infinitive Type</b>
2	<b>Internal Cause</b>	<b>Internal Cause</b>

TABLE 5.4 – Most significant factors influencing auxiliary selection as identified by CHAID and RF (Italian Semi-auxiliary verbs)

These two complementary statistical methods, CHAID and Random Forest, established a hierarchy of the most significant factors in the selection of auxiliaries for semi-auxiliary verbs, confirming the hypothesis and ensuring a comprehensive and accurate analysis of these particular verbs.

### 5.5 CHAID AND RANDOM FOREST FOR FRENCH VERBS

**CHAID** In contrast to Italian verbs, CHAID’s performance with French verbs is comparatively lower. The corpus analysis has shown that auxiliary selection in French is not governed by the same principles as in Italian. On the one hand, ‘have’ is more frequent; on the other hand, for one and the same verb, the selection of either ‘have’ or ‘be’ depends on the construction. Consequently, it becomes challenging to establish a definitive splitting node based on general criteria.

Nonetheless, Figure 5.5 indicates that the locative adverbial serves as this splitting node.

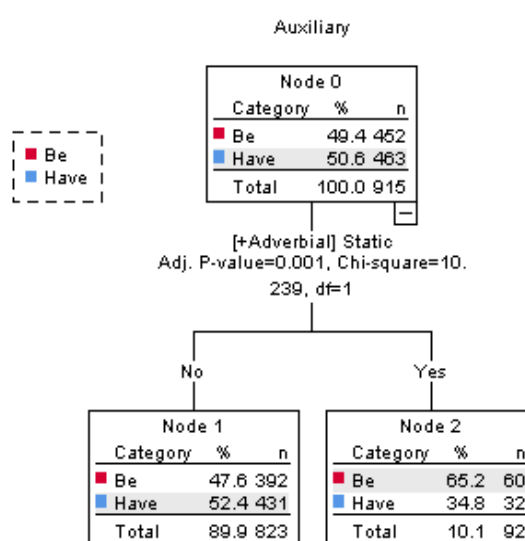


FIGURE 5.5 – Decision tree of the 10 French Verbs

Node 0 encompasses the entire corpus of the ten intransitive verbs analyzed in this study, comprising 990 sentences. It is divided into two sub-nodes based on what CHAID identifies as the most significant predictor ( $p$ -value < 0.001) of auxiliary selection, namely the locative adverbial indicating a stative location.

Node 2 exhibits the presence of the locative adverbial, predominantly featuring the verb ‘be’ (65.2%). This prevalence aligns with expectations, as corpus analysis has revealed a significant number of spatial motion verbs. These verbs typically align with ‘be’ when used in spatial motion constructions, resulting in a higher occurrence of ‘be’ with adverbials of spatial location.

Node 1 indicates the absence of the locative adverbial, with an almost equal distribution (50%) of ‘have’ and ‘be’. This equilibrium accurately reflects the general behavior of French intransitive verbs, which include various types (spatial motion, change of state, and verbs of appearance), when not differentiated based on their constructions.

For French verbs, the overall percentage of correctly classified cases is lower compared to Italian: the parameters assigned to the sentences in the sample enabled CHAID to accurately predict the choice of auxiliary in 53.7% of the cases, as shown in Table 5.4.

<b>Classification</b>			
	Predicted		
Observed	Be	Have	Percent Correct
Be	60	392	13,3%
Have	32	431	93,1%
Overall Percentage	10,1%	89,9%	53,7%

TABLE 5.5 – Classification table for CHAID model (French verbs)

Although the adverbial indicating static spatial location is statistically significant, the overall prediction accuracy is 53.7%. This percentage reflects the ability of the 16 identified parameters to correctly classify sentences and predict the auxiliary, which is challenging for French verbs.

This challenge was anticipated by the corpus analysis, which has shown that auxiliary selection is construction-dependent, not only for the selected sample of French intransitive verbs as a whole, but even for individual verbal lexemes that offer various possible constructions. Since the global analysis of the selected French verbs does give barely significant results, I have opted to conduct an analysis focused on different groups of verbs and of verb constructions.

**RANDOM FOREST** The analysis of the 10 verbs conducted with Random Forest (RF) corroborates the findings of the CHAID and my corpus analysis, by highlighting the fact that adverbials that provide a clue to a spatial movement construction tend to favor ‘be’, not only adverbials indicating a static location but even more so adverbials marking a direction with an endpoint.

Conversely, adverbs that emphasize the ongoing process, namely manner adverbs, but also adverbs conveying atelic aspect, favor the auxiliary ‘have’. The top group also contains adverbials of quantity, which frequently appear in quantitative evolution constructions.

Semantic factors relating to the subject such as internal cause and human trait follow in subsequent positions. As observed during the corpus analysis, these parameters may be correlated with certain types of verbal constructions, some of which select ‘be’ (e.g. spatial motion constructions), while others tend to select ‘have’ (e.g. quantitative evolution constructions).

Figure 5.6 presents the ranking results of the analysis conducted with Random Forest.

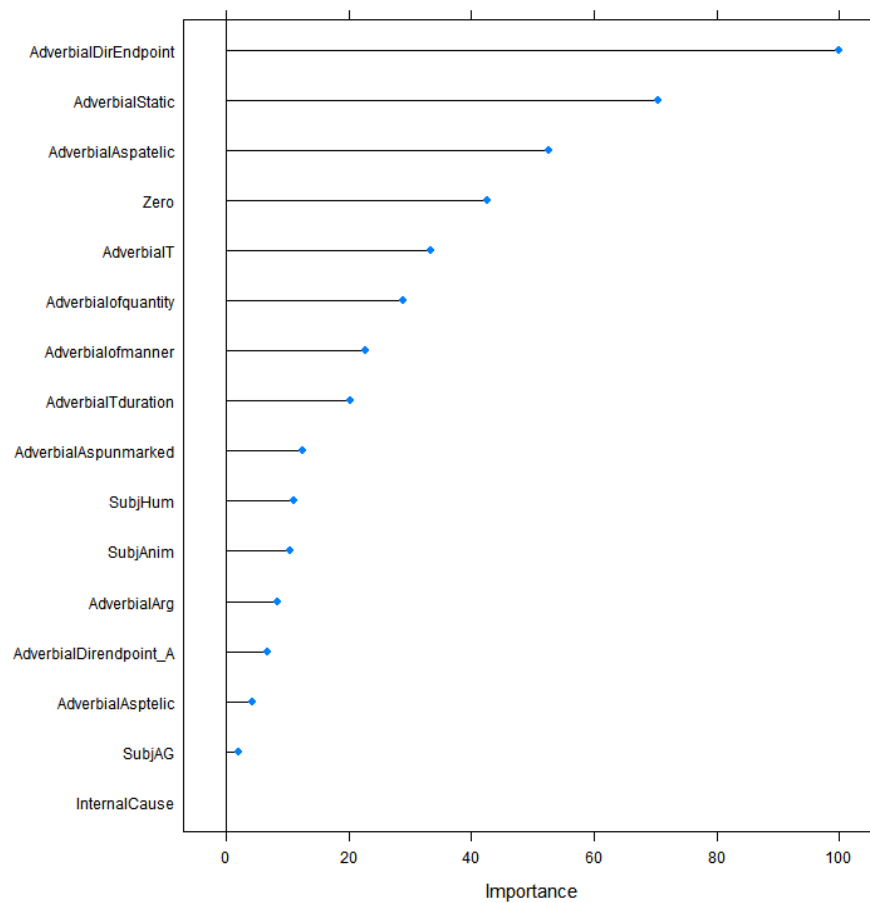


FIGURE 5.6 – Ranking of the sixteen parameters as revealed by RF (French verbs).

Figure 5.6 demonstrates that the outputs from random forest align with those from CHAID, though they are more precise. This alignment indicates that even with French verbs, the two statistical methods support the hypotheses established during corpus analysis.

Table 5.6 provides a summary of the most significant factors identified by both algorithms, CHAID and RF.

Most significant factors affecting auxiliary selection	
CHAID	RF
1 <b>Adverbial indicating static location</b>	<b>Locative adverbial indicating endpoint</b>
2	<b>Adverbial indicating static location</b>
3	Atelic aspectual adverbial
4	Adverbial of quantity

TABLE 5.6 – Most significant factors influencing auxiliary selection as identified by CHAID and RF (French verbs)

As previously noted in section 5.3 regarding the analysis with Italian verbs, there are slight differences in the outputs due to the distinct methodologies of CHAID and RF. However, despite these variances, the top positions remain consistent, with both methods consistently identifying the adverbials relating to spatial location as the most predominant factors influencing auxiliary distribution in French.

#### 5.5.1 CHAID AND RANDOM FOREST WITH *SPATIAL MOTION VERBS* (*passer, monter, baisser, descendre, diminuer*)

Given that French verbs can be classified into distinguishable categories with different behaviors, it seemed important to apply statistical analysis to the different groups of verbs, starting with the spatial motion verbs.

*Diminuer* ‘decrease’ has previously been considered a change-of-state verb. However, the corpus analysis showed that its behavior is very similar to *baisser*, as it almost exclusively presented the quantitative evolution construction (Chapter 4, Paragraph 4.2.3). Therefore, it seemed appropriate to include it within the group of verbs that effectively account for the quantitative evolution construction.

**CHAID** As demonstrated by the corpus analysis, verbs belonging to the spatial motion class show great sensitivity to the construction type. The construction type appears to be the primary

and most significant splitting factor, with a highly significant  $p$ -value of  $< 0.0001$ , as shown in Figure 5.7.

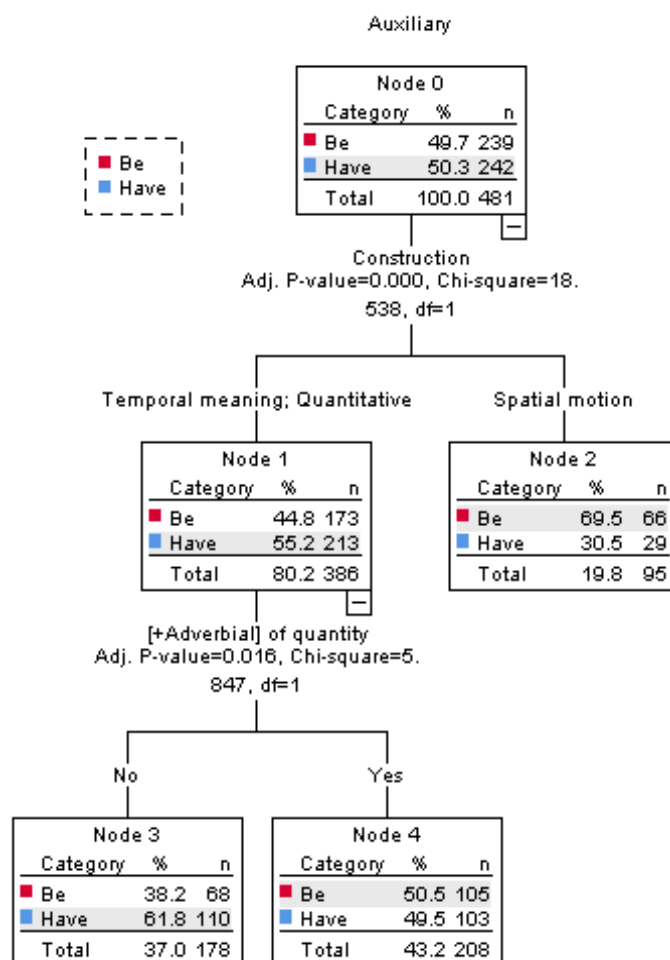


FIGURE 5.7 – Decision tree of the Spatial Motion French Verbs

Construction is the primary factor in determining the distribution of auxiliaries. With a spatial motion construction (Node 2), the most frequently selected auxiliary is ‘be’ (70%).

In contrast, when the construction is quantitative (or temporal, in the case of *passer*), depicted in Node 1, the distribution is more balanced, with a larger proportion of ‘have’ (55%) in comparison to ‘be’ (45%). This balance is likely due to the quantitative evolution construction’s potential to select ‘be’ with some of the analyzed verbs.

Node 3 reveals another relevant splitting factor, the adverbial of quantity, when the construction is quantitative or temporal. In the absence of the quantity adverbial, ‘have’ is the more frequently used auxiliary (61%). As shown by the corpus analysis, the absence of a

quantity adverbial does not imply the absence of the quantitative evolution construction (Node 3 includes temporal or quantitative evolution constructions without a quantity adverbial).

However, when the quantity adverbial is present, the percentages of auxiliary selection become more equal: instead of 60% ‘have’ and 40% ‘be’, Node 4 shows a 50% selection for both ‘have’ and ‘be’. This is because, in the corpus, some verbs in the quantitative evolution construction align with ‘be’ (e.g., *passer*), others with ‘have’ (e.g., *monter*, *descendre*), and some verbs present the quantitative evolution construction with both ‘have’ and ‘be’ (e.g., *baisser*, *diminuer*).

The overall percentage of correctly predicted auxiliary selection is lower than in Italian, but higher than in the analysis including all French verbs. As shown in Table 5.7, in nearly 60% of cases, the parameters assigned to the sentences in the sample enabled CHAID to accurately predict the choice of auxiliary for a given verb from the list.

Classification			
	Predicted		
Observed	Be	Have	Percent Correct
Be	171	68	71,5%
Have	132	110	45,5%
Overall Percentage	63,0%	37,0%	58,4%

TABLE 5.7 – Classification table for CHAID model (French Spatial motion verbs)

As depicted in Figure 5.7, in nearly 60% of cases, the parameters assigned to the sentences in our sample enabled CHAID to accurately predict the choice of auxiliary for a given verb from the list. This percentage might seem low (<70%), but this is because motion verbs are polysemic.

While spatial motion verbs used in a properly spatial motion construction predominantly select be ‘be’, these same verbs can use either ‘have’ or ‘be’ within the other constructions, particularly in quantitative evolution constructions. This overlap contributes to the lower prediction accuracy.

**RANDOM FOREST** The Random Forest statistical analysis of the five verbs confirms the findings from the previous CHAID and corpus analyses.

The hierarchy depicted in Figure 5.8 summarizes the most important factors influencing the distribution of the auxiliary.

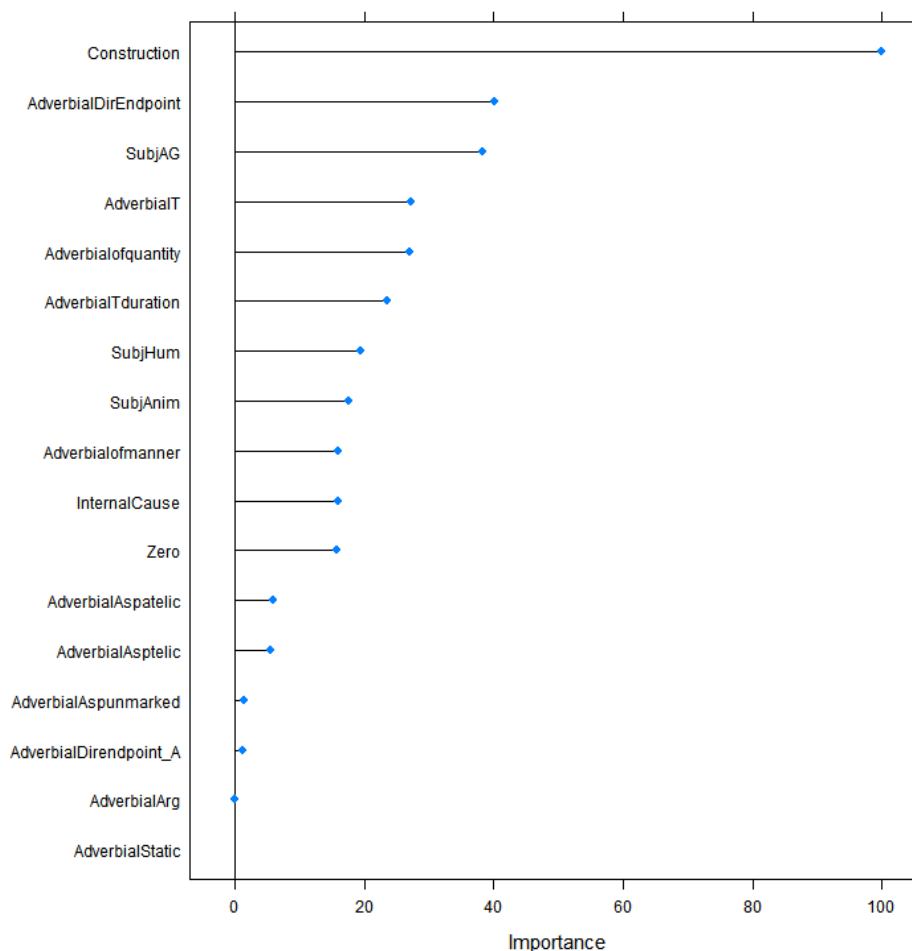


FIGURE 5.8 – Ranking of the sixteen parameters as revealed by RF (French Spatial motion verbs).

The type of construction turns out to be a decisive factor for the choice of the auxiliary. Other crucial but less important factors are adverbs indicating direction or an endpoint which are constitutive of spatial motion construction, as well as the agentivity of the subject. These three factors collectively correspond to the spatial motion construction, which typically selects ‘be’.

When the construction lacks agentivity and includes an adverbial of quantity, it translates into a temporal meaning or a quantitative evolution construction. In these cases, ‘have’ can prevail, or both ‘have’ and ‘be’ are accepted depending on the verb.

Table 5.8 provides a summary of the most significant factors identified by both algorithms, CHAID and Random Forest.

Most significant factors affecting auxiliary selection		
	<b>CHAID</b>	<b>RF</b>
1	<b>Construction</b>	<b>Construction</b>
2	<b>Adverbial of quantity</b>	Spatial location adverbials
3		Agentivity
4		<b>Adverbial of quantity</b>

TABLE 5.8 - Most significant factors influencing auxiliary selection as identified by CHAID and RF (Spatial motion French verbs)

By comparing the two statistical methods, the results are corroborated: the construction type is the most significant factor. The other two elements—locative adverbials, agentivity and adverbial of quantity—are largely correlated to the first element, the construction type.

#### 5.5.2 CHAID AND RANDOM FOREST WITH *CHANGE-OF-STATE VERBS* (*changer, mûrir, exploser*)

A second category of verbs analyzed consists of change-of-state verbs, namely *changer, mûrir*, and instances of *exploser* that convey the meaning of a process rather than a quantitative evolution construction.

These verbs are not distinguished on the basis of a specific construction, and they share more similarities with the peripheral verbs analyzed for Italian.

**CHAID** The primary statistically significant factor for change-of-state verbs has been revealed to be the semantic parameter of animacy, with a  $p$ -value  $< 0.0001$ , as illustrated in Figure 5.9.

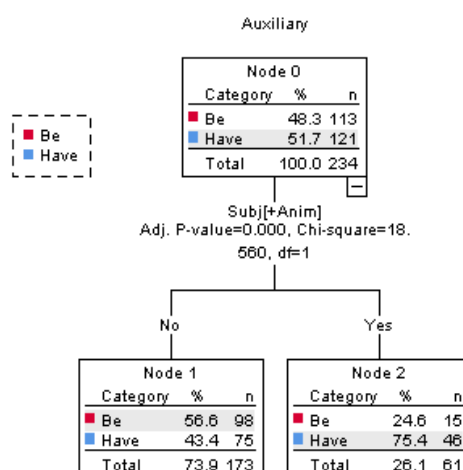


FIGURE 5.9 – Decision tree of the French change-of-state Verbs

Node 1 contains the relative frequency of ‘have’ and ‘be’ when the subject is inanimate. In this case, there is a slight preference for ‘be’, but the percentages of ‘have’ and ‘be’ selection are balanced (around 57% for ‘be’ vs. 43% for ‘have’).

In the presence of the animacy trait (Node 2), ‘have’ is highly prevalent, with 75% selection compared to 25% for ‘be’. Similar to Italian, the animacy (or human trait) significantly influences the selection of ‘have’ for French change-of-state verbs.

Even when restricted to change-of-state verbs, the overall percentage of correct auxiliary selection obtained by the CHAID analysis is lower than that for Italian verbs but higher than that for the general corpus of French verbs, as illustrated in Table 5.9.

Classification			
	Predicted		
Observed	Be	Have	Percent Correct
Be	98	15	86,7%
Have	75	46	38,0%
Overall Percentage	73,9%	26,1%	61,5%

TABLE 5.9 – Classification table for CHAID model (French change-of-state verbs)

The table shows that in 61.5% of cases, animacy is correctly predicted as a significant factor influencing the distribution of ‘have’ and ‘be’. This means that animacy is a key determinant in auxiliary selection for change-of-state verbs.

**RANDOM FOREST** The Random Forest statistical analysis of the three change-of-state verbs confirms the findings from the previous CHAID and corpus analyses. Similar to Italian verbs, semantic features characterizing the subject play a key role in auxiliary distribution (Figure 5.10).

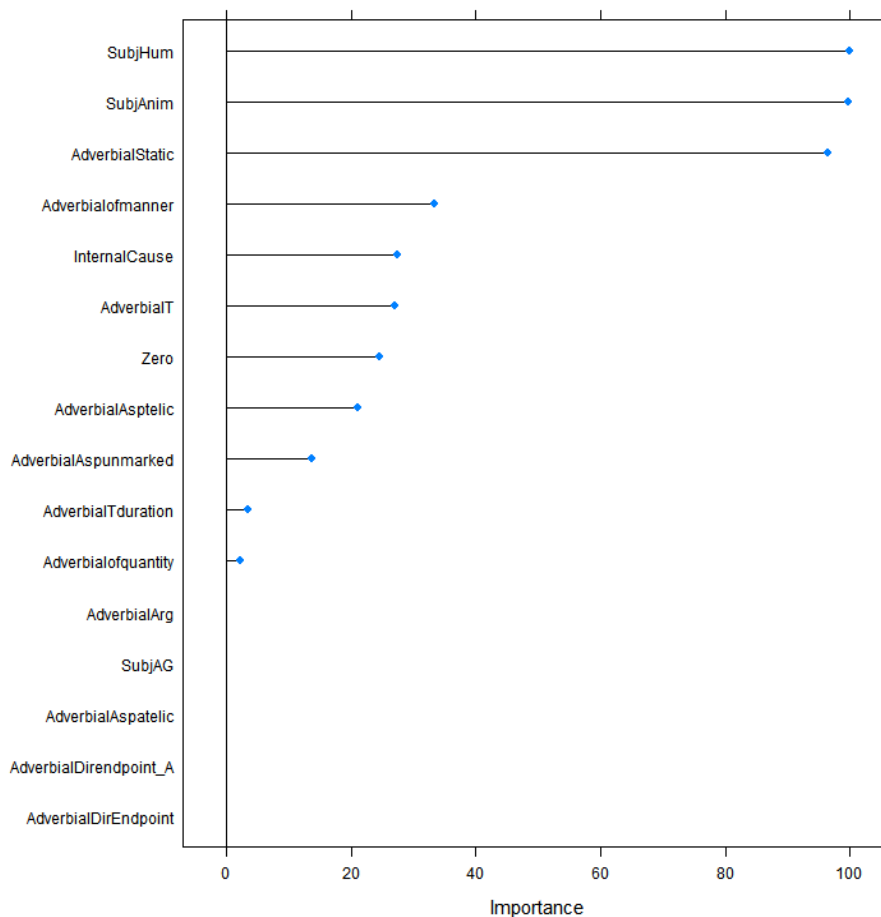


FIGURE 5.10 – Ranking of the sixteen parameters as revealed by RF (French change-of-state verbs).

The key relevant factors are the semantic traits of the subject, namely its human or animate nature, which correlate with the selection of ‘have’.

Another relevant factor is the locative adverbial, though it is less important due to the nature of the verbs. The analysis has demonstrated that some factors may appear in the context of the sentence despite not having an impact on auxiliary selection (e.g., *les liens [...] dans la ville étaient explosés* ‘The links [...] in the city were exploded’). For this reason, the analysis through both methods proved crucial in identifying the factors that are truly relevant for the choice of the auxiliary.

When comparing the two statistical analyses, the relevant factors are confirmed (Table 5.10):

Most significant factors affecting auxiliary selection		
	CHAID	RF
1	Animacy	Human Trait
2		Animacy
3		Locative adverbial

TABLE 5.10 – Most significant factors influencing auxiliary selection as identified by CHAID and RF (French change-of-state verbs)

### 5.5.3 CHAID AND RANDOM FOREST WITH VERBS OF APPEARANCE (*paraître, apparaître*)

The third and final group of verbs in the analysis consists of verbs of appearance. During the corpus analysis, these verbs were found to be primarily distinguished on the basis of their construction.

**CHAID** The primary statistically significant factor for verbs of appearance has been revealed the type of construction, with a  $p$ -value  $< 0.0001$ , as illustrated in Figure 5.11.

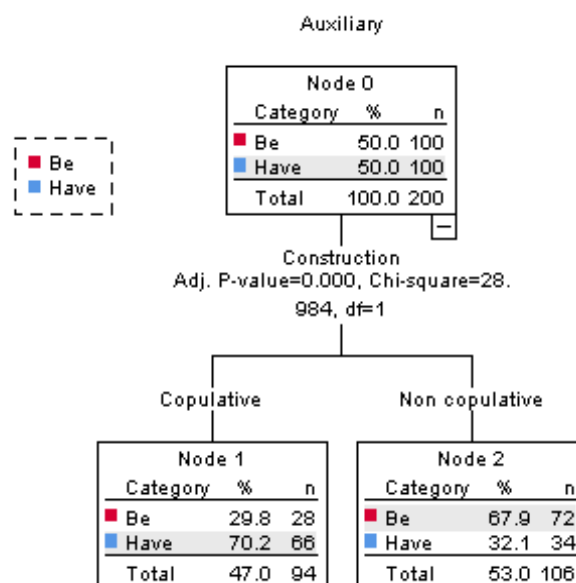


FIGURE 5.11 – Decision tree of the French Verbs of Appearance

When the construction is copulative (Node 1), the preferred auxiliary is ‘have’, with 70% selecting ‘have’ and 30% selecting ‘be’.

When the construction is non-copulative (Node 2), the situation is reversed: the favored auxiliary is ‘be’, with 68% selecting ‘be’ and 32% selecting ‘have’.

The outputs of the CHAID analysis corroborate the findings from the corpus analysis, and the overall accuracy is even higher than for the previous group: in almost 70%, of cases, the type of construction allows to predict whether the auxiliary ‘have’ or ‘be’ is used, as demonstrated in Table 5.11.

<b>Classification</b>			
Observed	Predicted		
	Be	Have	Percent Correct
Be	72	28	72,0%
Have	34	66	66,0%
Overall Percentage	53,0%	47,0%	69,0%

TABLE 5.11 – Classification table for CHAID model (French Verbs of appearance)

**RANDOM FOREST** The hierarchy resulting from Random Forest confirms the analysis conducted via CHAID, indicating the type of construction as the most relevant factor, as shown in Figure 5.12.

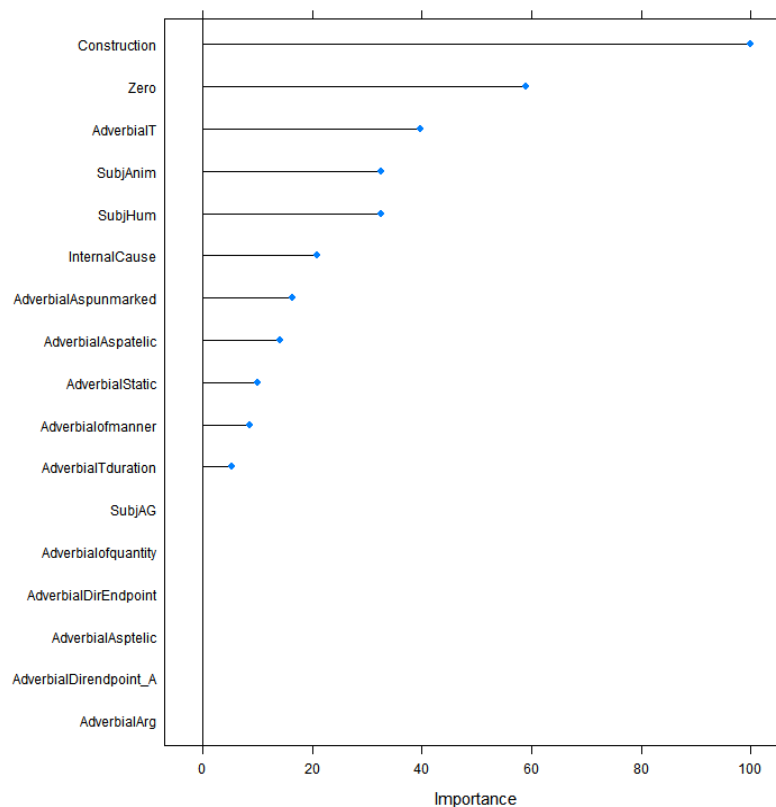


FIGURE 5.12 – Ranking of the sixteen parameters as revealed by RF (French Verbs of appearance).

The type of construction emerges as the most influential factor affecting the distribution of the auxiliary, as indicated by the length of the line in the diagram. Other factors listed afterward were found to be less relevant and rather associated with the construction in which they were embedded.

In conclusion, both statistical methods of analysis yielded consistent results for verbs of appearance, highlighting the type of construction as the primary determinant in the distribution of the auxiliary, as illustrated in Table 5.12.

Most significant factors affecting auxiliary selection	
CHAID	RF
1	<b>Construction</b>
2	Zero
3	Temporal Adverbial

TABLE 5.12 – Most significant factors influencing auxiliary selection as identified by CHAID and RF (French Verbs of appearance)

## 5.6 CONCLUDING REMARKS

The statistical analysis conducted by means of two methods, CHAID and Random Forest, has been of a crucial importance in my research. Not only did it corroborate the hypotheses build on the basis of the corpus analysis by providing evidence of the reliability (in terms of statistical significance) and reproducibility of findings, but it also allowed us to assign weight to the different parameters or features based on their importance in predicting auxiliary selection in both languages.

The reliability of the parameters in determining the choice of auxiliary proves to be higher for Italian than for French. For Italian intransitive main verbs, the parameters conceived can explain over eighty percent of auxiliary alternation for the verbs analyzed, which is a remarkable result.

For Italian semi-auxiliaries, the results show that the parameters used (infinitive type and semantic features) can correctly predict the auxiliary used 77% of the time.

For French, accuracy is notably lower, particularly when the entire sample of verbs is analyzed collectively. To achieve more precise results, it was necessary to partition the sample into subclasses. As previously noted, the difference in predictability percentages between Italian and French—80% versus 53.7%, respectively—can be largely attributed to differences in the way these languages handle auxiliary alternation, especially when verb classes are not

distinguished. Italian verbs are more consistently than French verbs subject to alternation in function of the central parameter of internal cause.

Paradoxically, French verbs can more readily take both ‘have’ and ‘be’ auxiliaries, as demonstrated by examples like *Il a monté* and *Il est monté*. At first glance, this might suggest that auxiliary alternation in French is more arbitrary. However, a more detailed analysis shows that auxiliary selection in French is influenced by the verb's syntactic construction: the same verb can function differently based on its syntactic context, affecting its auxiliary selection and overall meaning. For example, the verb *monter* in a spatial construction (e.g., *Il est monté au grenier* ‘He went up to the attic’) functions differently from *monter* in a quantitative evolution construction, as in *Les prix ont monté de 3 %* ‘Prices have risen by 3%’.

Table 5.13 offers an overview of the most significant factors highlighted by the statistical analysis using both CHAID and Random Forest for each verb class.

VERB CLASS	FACTORS
Italian Verbs (General)	1.Internal Cause 2.Human Trait
Italian Verbs (Semi-auxiliaries)	1.Type of Infinitive 2.Internal Cause
French Verbs (General)	1.Locative Adverbials indicating static location
French Verbs (Spatial motion)	1.Construction
French Verbs (Change-of-state)	1.Animacy 2.Human Trait
French Verbs (Appearance)	1.Construction

TABLE 5.13 – Hierarchy of the most significant factors identified by CHAID and RF for each class of verbs.

As Table 5.13 illustrates, the most significant factors for Italian main verbs are internal cause and human trait, both semantic factors characterizing the subject.

For Italian semi-auxiliaries, the factors accounting for the choice of auxiliary are auxiliary selection properties of the infinitive verb and, corollarily, the nature of the subject (agentive, internal cause, or non-agentive).

For French change-of-state verbs, the crucial factors are also semantic factors characterizing the subject, viz. animacy and human trait.

Hence, both Italian peripheral verbs and a subset of French peripheral verbs, *i.e.* change-of-state verbs include verbs where the subject can often be interpreted as both the agent and patient

of the same action. Thus, the agentivity/internal cause of the subject and their human or animacy trait play a significant role.

However, other French verbs exhibit different behavior from the Italian verbs, showing locative adverbials as well as quantity adverbials as particularly significant. For spatial motion verbs and verbs of appearance, the syntactic construction emerged as the most influential factor in auxiliary selection. Spatial motion verbs distinguish between spatial motion constructions and quantitative evolution constructions, while verbs of appearance differentiate between copulative and non-copulative constructions.

Thus, while semantic parameters related to the nature of the subject are paramount for Italian peripheral verbs and French change-of-state verbs, other French verbs rely more heavily on syntactic construction. In the latter case, adverbials play a role, insofar as they are a constitutive element of one construction or another or, more occasionally, provide aspectual information.

The application of statistical tools has been instrumental in highlighting these tendencies with precise data for both Italian and French peripheral verbs. The sample of verbs analyzed differ between the two languages, but the analysis revealed this is due to the fact that auxiliary alternation doesn't involve exactly the same verbs in both languages. A case in point are spatial motion verbs, which are 'be' selecting in Italian and which allow for 'be' or 'have' depending on the chosen construction in French.

## CHAPTER 6: COMPARATIVE ANALYSIS OF AUXILIARY SELECTION IN ITALIAN AND FRENCH

This chapter presents a comparative examination of auxiliary distribution in Italian and French, aiming to elucidate nuanced distinctions and common trends in these languages. The analysis focuses on how semantic factors and syntactic constructions influence the selection of auxiliaries.

Section 6.1 explores the main factors influencing auxiliary distribution in Italian. Section 6.2 focuses on the outputs of the analysis in French. Finally, Section 6.3 delineates the similarities and differences between the two languages.

### 6.1 AUXILIARY DISTRIBUTION IN ITALIAN: THE OUTPUTS OF CORPUS ANALYSIS AND STATISTICAL ANALYSIS

Qualitative and quantitative corpus analysis distinguished for Italian two categories of verbs, viz. verbs used as full verbs and semi-auxiliary verbs. Since specific parameters of semi-auxiliaries with respect to auxiliary selection have been detected only for Italian verbs in the peripheral verbs analyzed in this comparative study, this section will focus only on the results of full verbs.

*Qualitative Corpus Analysis* The corpus analysis underscored the importance of two semantic factors influencing the distribution of auxiliaries: [ $\pm$  HUMAN TRAIT] and the nature of the subject in terms of agentivity, non-agentivity, or internal causation. The presence of human subjects correlates with the auxiliary ‘have’ (Chapter 3), as does the presence of agentive subjects. Conversely, the absence of human traits and agentivity correlates with the auxiliary ‘be’.

In comparison with the state of the art, the present investigation highlights not only the feature [+HUMAN] as a sufficient condition to trigger the choice of the auxiliary ‘have’, but also, and more importantly, the central role of internal causation in auxiliary selection.

Indeed, the concept of internal causation, where the subject can simultaneously function as both the agent and the patient of the same action, plays a significant role in determining whether ‘have’ or ‘be’ is used as the auxiliary. This nuanced understanding highlights how semantic

factors such as internal causation contribute to the variability in auxiliary selection across different linguistic contexts.

These findings underscore the complexity of auxiliary usage in relation to semantic roles within sentences.

*Statistical Analysis* The findings from the corpus analysis have been validated through the use of statistical methods, which provided precise numbers and percentages. This statistical analysis enabled us to quantify the distribution of ‘have’ and ‘be’ auxiliaries based on various factors and to evaluate the significance of each factor.

Specifically, agentive subjects exhibit a 100% correlation with ‘have’, while non-agentive subjects show a 92,2% correlation with ‘be’. Subjects having the role of internal cause without being agentive display a distribution of 37,9% with ‘be’ and 62,1% with ‘have’, illustrating a pattern of alternation between the two auxiliaries.

It proved to be important to isolate and further analyze other factors that influence the choice of auxiliary when the subject is neither fully agentive nor devoid of agentivity, as this scenario alone represents 44% of the sample of 1146 analyzed occurrences. It is in this configuration that the [+HUMAN] characteristic of the subject, not mentioned in earlier studies, has proven highly effective, as its presence triggers a 90% of ‘have’ selection.

Hence, using CHAID and Random Forest methods, the statistical analysis allowed us to establish a hierarchy of the most influential factors. Primary among these factors is the nature of the subject in terms of agentivity, non-agentivity, or internal causation, followed by the human trait, the latter feature being especially pertinent for predicting the selection of the auxiliary when the subject corresponds to a non-agentive internal cause.

## 6.2 AUXILIARY DISTRIBUTION IN FRENCH: THE OUTPUTS OF CORPUS ANALYSIS AND STATISTICAL ANALYSIS

French verbs exhibit a distinct pattern in auxiliary selection compared to Italian verbs. Firstly, the procedure for selecting verbs with auxiliary alternation resulted in a sample that was quite different. Secondly, unlike Italian verbs, French verbs are notably sensitive to the syntactic constructions in which they appear.

**Qualitative Corpus Analysis** The corpus analysis of French verbs highlighted that these verbs can be divided into three different categories: spatial motion verbs, change-of-state verbs, and verbs of appearance. Each category demonstrated different significant factors.

Spatial motion verbs, when used in a spatial motion construction, select the auxiliary ‘be’. When used in a quantitative evolution construction, the auxiliary is ‘have’. The verb *passer* is particular insofar it selects ‘be’ in the quantitative evolution construction, and additionally presents a temporal construction, with ‘have’).

For verbs of appearance, if the construction is copulative, the auxiliary is ‘have’; otherwise, it is ‘be’ (at least for *paraître*, since *apparaître* allows for the use of both auxiliaries in both constructions).

The human trait emerged as another significant factor in the choice between ‘have’ or ‘be’. However, for spatial motion verbs and verbs of appearance, this feature is rather contingent on the type of construction. For change-of-state verbs, the presence of a [+ HUMAN TRAIT] is plainly relevant and strongly correlated with the use of ‘have’. This similarity makes change-of-state verbs the category most similar to Italian verbs allowing both auxiliaries (especially *changer* and *mûrir*).

The [± AGENTIVITY] factor displayed a comparable trend. Spatial motion verbs and verbs of appearance showed a correlation between agentivity and construction type. In contrast, for change-of-state verbs, agentivity correlated with the use of ‘have’, mirroring patterns observed in Italian verbs.

In summary, while construction type primarily determines auxiliary selection for spatial motion verbs and verbs of appearance, the human trait, agentivity and internal cause play critical roles in change-of-state verbs, aligning with findings from the analysis of Italian verbs.

**Statistical Corpus Analysis** The findings from the corpus analysis have been validated through statistical methods, including for French verbs, particularly when categorized into spatial motion verbs, change-of-state verbs, and verbs of appearance.

In spatial motion verbs and verbs of appearance, the type of construction emerged as the most influential factor. ‘Be’ is used in 70% of spatial motion constructions.

For verbs of appearance, ‘have’ is preferred in copulative constructions 70% of the time, compared to 30% for ‘be’. In non-copulative constructions, the trend is reversed.

Change-of-state verbs were notably influenced by animacy as the primary factor, with animate subjects showing a 75% preference for ‘have’. These verbs exhibit behaviors closely

aligned with Italian verbs (especially *changer* and *mûrir*), emphasizing the impact of semantic parameters associated with the subject on auxiliary selection.

### 6.3 A COMPARATIVE ANALYSIS OF AUXILIARY DISTRIBUTION IN ITALIAN AND FRENCH

The combined use of qualitative corpus analysis and statistical methods applied on the same data has been crucial in substantiating hypotheses regarding the factors influencing the distribution of ‘have’ and ‘be’. Employing both methods has provided robust evidence to support these findings.

The state of the art primarily highlights that auxiliary alternation is a declining phenomenon in French, with the use of ‘have’ expanding at the expense of ‘be’. Legendre and Sorace (2003: 11) argue that “many verb classes in French have already transitioned to using *avoir*, resulting in less variation and gradience compared to Italian.” In other words, in their view, on the scale established within the framework of the ASH theory, French and Italian do not occupy the same position. Italian verbs exhibit greater gradience and variability in auxiliary distribution, whereas French verbs may employ both auxiliaries due to the extended usage of ‘have’ into contexts where ‘be’ was traditionally used.

Firstly, to validate this hypothesis, a diachronic study of changes in auxiliary selection throughout the history of French is needed.

Secondly, my study highlighted that the difference between Italian and French concerns not only the number of verbs that allow ‘be’ or both auxiliaries, but also, and more importantly, the principles governing the choice of the auxiliary. For example, the choice between ‘have’ and ‘be’ with spatial motion verbs depends on the construction used, which can be either a spatial movement construction or a quantitative change construction, as discussed in section 6.2. This means that agentivity, highlighted as one of the major parameters for explaining auxiliary selection in the ASH scale, does not have the same explanatory power for French as it does for Italian<sup>26</sup>.

However, French and Italian verbs display analogous patterns in the case of change-of-state verbs. French change-of-state verbs resemble Italian verbs that exhibit auxiliary alternation, where semantic factors such as human or animate traits and the nature of the subject as agentive, non-agentive, or involving internal causation play crucial roles. At least for this type of verbs,

---

<sup>26</sup> A confirmation is provided by Heidinger’s (2015) diachronic study on the auxiliary selection of French *monter* from the 16th to the 20th c.

it is possible to identify common underlying principles at work in auxiliary selection for both languages.

Table 6.1 summarizes the key factors influencing verbs in French and Italian, providing a structured overview of their respective patterns of auxiliary usage.

ITALIAN MAIN VERBS	FRENCH CHANGE-OF-STATE VERBS	FRENCH SPATIAL MOTION VERBS VERBS OF APPEARANCE
1.Agentivity and Internal Cause 2.Human Trait	1.Human trait 2.Agentivity and Internal Cause	1.Construction

TABLE 6.1 – Hierarchy of factors influencing auxiliary distribution

Table 6.1 illustrates the hierarchy of factors relevant to the distribution of auxiliaries in French and Italian. Corpus analysis initially helped identify the factors influencing auxiliary selection, while statistical analysis established the hierarchy of these factors, determining their relative importance.

In Italian, the primary factors identified were the nature of the subject as agentive, non-agentive, or involving internal causation, along with the human trait of the subject. Internal cause emerged as a significant factor contributing to variance in auxiliary selection. This phenomenon characterizes the subject as both agent and, at other times, patient of the action, leading to challenges in definitively determining the selected auxiliary.

When the subject exhibits internal cause, the human trait becomes the second influential factor. Specifically, if the subject is agentive, the auxiliary ‘have’ is typically chosen; if non-agentive, ‘be’ is often preferred, with minimal impact from the human trait. However, in cases where the subject is both human and an internal cause, the human trait significantly influences auxiliary selection. ‘Have’ is selected in 90% of occurrences under these conditions (refer to Section 5.3 in Chapter 5 for more details).

For French, on the other hand, a subcategorization of peripheral verbs into several classes has proven necessary, as no single parameter effectively predicts the choice of auxiliary for the peripheral verbs as a whole. For some of these sub-categories, such as spatial motion verbs and verbs of appearance, the type of construction is the statistically significant factor. These verbs exhibit varied behavior depending on the construction: spatial motion verbs, for instance, use ‘be’ in spatial motion constructions and ‘have’ in quantitative evolution constructions (except

*passer*). Other factors, such as locative adverbials for spatial motion verbs, are pertinent only to the extent that they help shape the construction.

The remaining class of verbs, the change-of-state verbs, behaves more similarly to Italian verbs with double auxiliary selection (especially the verbs *changer* and *mûrir*). These verbs are relatively uniform and emphasize agentivity and internal causation as significant factors, second only to the human trait. When the subject is animate (and thus human), ‘have’ is typically the auxiliary of choice. This category of verbs maintains consistency across various constructions, reflecting similar characteristics as observed in Italian. Here, the distinction between animate/human subjects versus inanimate/non-human subjects, and the potential for the subject to act as both agent and patient, remains pertinent in explaining the observed variation in auxiliary usage.

In conclusion, both corpus analysis through annotation and statistical methods were crucial in uncovering the factors determining auxiliary selection and the conditions allowing for variance and the possibility of using both auxiliaries.

In Italian, this variability primarily stems from semantic parameters associated to the subject, whereas in French, it crucially depends on the verb’s syntactic construction and only secondarily on semantic factors.

Corpus analysis played a pivotal role in revealing patterns of auxiliary usage and in understanding the similarities and differences between French and Italian, while statistical analysis provided an ordered hierarchy of influential factors.

## CONCLUSION

Auxiliary selection is a crucial topic in linguistic theory, as it reveals intricate relationships between syntax, semantics, and morphology. This thesis sets out to investigate auxiliary selection in two Romance languages, Italian and French. The primary objective was to understand the factors influencing the choice between the auxiliaries ‘be’ and ‘have’ in the formation of the perfect tense. By examining both qualitative and quantitative data, this study sought to answer key questions: What are the underlying reasons for the auxiliary split? How do semantic and syntactic factors contribute to this choice? And how consistent are these patterns across different languages?

### SUMMARY OF KEY FINDINGS

From a theoretical perspective, this study offers a historical overview of various theories, from the unaccusative hypothesis (Perlmutter 1979) to more nuanced frameworks like Sorace’s Auxiliary Selection Hierarchy (2000). It underscores the complexity and multifaceted nature of auxiliary selection, laying the foundation for subsequent analyses. While substantial research has explored the parameters influencing auxiliary selection — particularly in Italian — this study aims to contribute novel insights in several key areas.

Firstly, it delves into a relatively unexplored area of this research: the so-called peripheral verbs (Sorace 2000: 860), which can take either ‘have’ or ‘be’ in the perfect tense. For Italian verbs, our analysis innovates by demonstrating that the parameter of agentivity is insufficient for accounting for ‘have’ selection, since it accounts for only 256 perfect tenses with ‘have’ in our sample of 600 occurrences, a success rate of 43%. For this reason, the concept of ‘internal cause’ has been put forward. This concept, which bridges the agent and patient roles, provides a more nuanced framework for understanding auxiliary alternation.

Our study has also demonstrated the relevance of the [+HUMAN] feature, which is barely highlighted in previous studies. The combination of the internal cause feature and [+HUMAN] feature allows for a more accurate definition of the rules that determine the use of the auxiliary ‘have’ rather than ‘be’.

Another significant finding with respect to auxiliary selection in Italian is the specific behavior with respect to auxiliary selection of verbs like *iniziare*, *cominciare*, *continuare*, and

*finire*, when they combine with an infinitive verb form. This phenomenon has been previously discussed under the broader category of ‘Restructuring’ (Rizzi 1976a) but has not specifically analyzed in works dedicated to split intransitivity. Instead of ‘Restructuring’, which consists in a transformation of a verbal construction composed of two verbal predicates, namely an inflected verb and an infinitive, into a single verbal predicate, another view is taken in the present study.

In our perspective, the auxiliary alternation of four Italian verbs—*iniziare*, *cominciare*, *continuare*, and *finire*—when they combine with an infinitive verb form, can be accounted for by their status of semi-auxiliary, which leads them to convey certain grammatical markers of the infinitive verb, acting as the main verb.

Consequently, the auxiliary selection properties of the verb in the infinitive form play a crucial role in influencing the auxiliary selection of the semi-auxiliary verb. Specifically, when the infinitive is a transitive verb, the selected auxiliary will be ‘have’, whereas for intransitive verbs, ‘have’ selecting verbs favor ‘have’ for the semi-auxiliary verb, whereas ‘be’ selecting verbs allow ‘be’ for the semi-auxiliary verb.

Additionally, for infinitive verbs preferring ‘be’, or allowing both ‘have’ and ‘be’, the nature of the subject (human or non-human; agentive, non-agentive, or internal cause) emerges as influential. Notably, *finire* behaves differently from the other three semi-auxiliaries due to its compatibility with various prepositions, influencing its auxiliary choice.

With respect to French, our study reveals a previously unexplored aspect: the dependency of auxiliary choice on the construction type. Three categories of verbs exhibit auxiliary alternation: spatial motion verbs, change-of-state verbs, and verbs of appearance. For spatial motion and appearance verbs, auxiliary alternation is highly sensitive to construction type. As a result, one and the same verb can appear in different construction types and select an auxiliary depending on the construction in which it is embedded. Conversely, different verbs can share the same syntactic construction type and, as a consequence, exhibit similar auxiliary selection properties.

A case in point is *diminuer* ‘reduce’/ ‘decrease’, initially identified as change-of-state verb, but which in our data primarily functions in the quantitative evolution construction, and as such shares the properties of spatial motion verbs like *monter* ‘go up’ and *descendre* ‘go down’ when they are embedded in the quantitative change construction, by systematically featuring the verb ‘have’.

Whereas the major parameter for auxiliary selection in French is the construction type, the specific category of change-of-state verbs in French resemble Italian verbs allowing both auxiliaries, with semantic factors like internal cause/agentivity and human trait being pivotal. Again, the emphasis on the internal cause factor, highlighting the subject's dual role as agent and patient, is innovative compared to previous studies primarily focused on intransitive verbs (Mateu 2009) without specifically addressing peripheral verbs. It explains why both 'have' and 'be' may be chosen, and how semantic nuances may in a subtle way impact auxiliary selection. Even Sorace (2000: 862) considered the concept of internal cause, albeit in relation to the event structure of predicates and process verbs, but it has not previously been explicitly proposed as a determinant for the selection of a specific auxiliary<sup>27</sup>.

As to the French semi-auxiliaries of aspect like *commencer* 'begin', *continuer* 'continue', *cesser*, *finir* 'stop', – unlike their Italian counterparts in Italian – they are not peripheral verbs since they invariably select 'have' and therefore they have not been considered in this study.

The above considerations lead us to another original contribution, which concerns the comparison between French and Italian, two Romance languages that exhibit auxiliary alternation between 'have' and 'be' in the perfect tense. Previous studies comparing French and Italian have primarily observed that French, compared to Italian, is in the process of losing auxiliary alternation and expanding the use of 'have' as a default auxiliary. As Legendre and Sorace (2003: 11) put it: "many verb classes in French have already transitioned to using *avoir*, resulting in less variation and gradience compared to Italian". This study devoted to peripheral verbs has the merit of showing that French differs from Italian not only in the broader use of 'have', *i.e.* in combination with more verbs or verb classes, but also in the underlying principles governing auxiliary alternation.

These differences between the two languages should not obscure the fact that the status of 'have' as the default auxiliary is not confined to French. Regarding Italian aspectual semi-auxiliaries, this study has shown that 'be' is recessive. Indeed, when combined with an infinitive verb that selects the auxiliary 'have' in the perfect tense, the semi-auxiliary will select 'have'. However, when combined with an infinitive verb that allows both auxiliaries or even is

---

<sup>27</sup> Sorace mentions internal causation two times at page 862: (1) "The notion of causation separates monadic (e.g. *arrive*) from dyadic verbs of transition (e.g. *break*). Internal causation (which subsumes the notion of agentivity) is only of secondary relevance for monadic verbs of transition. (2) "There is an asymmetry between agentivity and internal causation: all process verbs are internally caused in the sense that they are brought about by a causer, but they vary systematically in the extent to which the causer is an intentional agent."

exclusively ‘be’ selecting, the semi-auxiliary will not necessarily select ‘be’, but ‘have’ may also appear.

In conclusion, this study delineates both the shared patterns between the two languages and the areas of dissimilarity, demonstrating that even genetically related languages do not share identical principles of auxiliary selection.

## **METHODOLOGICAL STRENGTHS OF THIS RESEARCH**

Our study was designed according to a rigorous methodology, encompassing data selection, annotation, and analysis. The decision to focus on Italian and French, languages known for their auxiliary variability, was pivotal for the study’s depth and scope in order to test the cross-linguistic validity of the principles of auxiliary selection. The sentences, which constitute the empirical basis of this study, were accurately retrieved using a specific query formula using Corpus Query Language (CQL) and according to explicit criteria, which guarantees the reproducibility of this study. By choosing the corpora *ItTenTen* (16), and *FrTenTen* (17), available via *SketchEngine*, a sufficient amount of data was ensured, comparable in terms of register for both languages. All sentences, both in French and Italian, have been systematically annotated for sixteen parameters that emerged during the analysis as potentially impacting the choice of auxiliary.

Crucial in our methodological approach is the choice of combining qualitative corpus analysis with statistical analysis of the same corpus data. This approach facilitated a thorough examination of auxiliary selection patterns and the identification of relevant parameters. By employing both methods, it was possible to not only identify the parameters explaining auxiliary selection but also to establish a hierarchy of these parameters by measuring their relative weight.

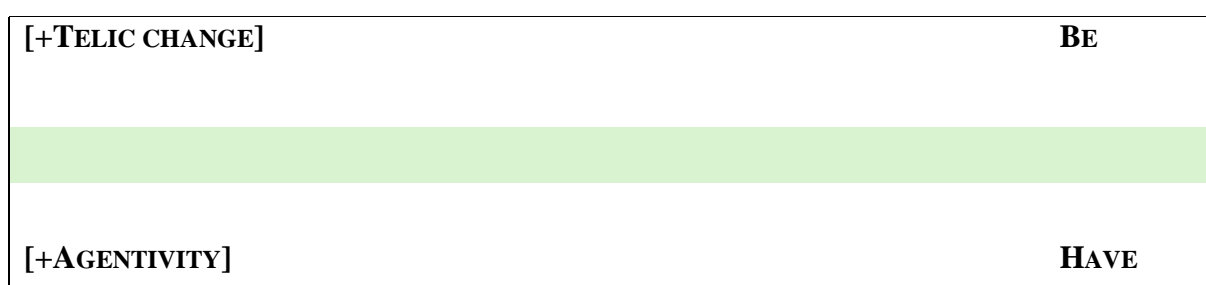
CHAID provides a clear view of the effectiveness of major parameters by numerically evaluating and ranking them, identifying the most significant factors influencing the outcome. In contrast, Random Forest offers a more detailed view, including parameters with lesser impact, by assessing the significance of all variables and exploring complex interactions. The complementarity of these tools lies in CHAID’s ability to highlight major factors while Random Forest provides depth and granularity, ensuring a comprehensive analysis that captures both high-level trends and detailed insights.

The study’s methodology, combining qualitative and quantitative analyses, provides a model for future research in linguistic studies. It demonstrates the value of an integrative approach in

unraveling complex linguistic phenomena, enabling a more comprehensive and nuanced analysis.

### **SORACE’S (2000) AUXILIARY SELECTION HIERARCHY: NEW INSIGHTS**

This conclusion gives us the opportunity to revisit Sorace’s (2000) Auxiliary Selection Hierarchy, structured as a scale with the feature of telicity correlated to ‘be’ selection at the top of the scale and the feature of agentivity at the bottom of the scale correlated to ‘have’ selection.



**FIGURE 1**– Illustration of the main concepts in Sorace’s Auxiliary Selection Hierarchy, highlighting telicity and agentivity in relation to auxiliary selection.

Various factors, such as telicity, agentivity, affectedness, and density, are considered relevant. However, when it comes to establishing the hierarchy, two key notions stand out: telicity and agentivity (2000: 861-862).

As demonstrated in the state of the art, the Auxiliary Selection Hierarchy model raises several issues, three of which will be highlighted here.

- The first issue pertains to the model’s cross-linguistic validity. The ASH is argued to be universal in that, for any given language, auxiliary selection rules are conceived in terms of a cut-off point on the ASH, distinguishing verbs that select auxiliary ‘be’ from those that select auxiliary ‘have’. While the location of this cut-off point cannot be identical across all languages (as this would result in all languages having the same system of auxiliary selection), language-specific variation is constrained by the ASH, in that it can only involve differences in the placement of the cut-off point along the hierarchy.
- The second issue is the lack of attention for so-called peripheral verbs, which allow for both auxiliaries ‘have’ and ‘be’.
- The third issue concerns the conception of the model as a scale with telicity and agentivity at its two ends. On the one hand, these features are neither antonymous nor complementary

and can therefore hardly be conceived as the extreme poles of a scale. On the other hand, they do not have the same impact on auxiliary selection. The example *Pierre est arrivé* shows that despite an agentive subject, the feature of telicity leads to the selection of ‘be’. Conversely, there are many cases such as *La temperature a monté*, where the chosen auxiliary is ‘have’ even when the subject is non-agentive. Particularly concerning the third issue mentioned above, the concept of agentivity proves overly simplistic for addressing the nuances of auxiliary selection. We propose replacing the binary opposition [ $\pm$ AGENTIVE] with a more refined and gradual framework that incorporates the features [ $\pm$ INTERNAL CAUSE] and [ $\pm$ HUMAN].

Based on the analysis of the theoretical approach and the statistical methods, we expanded Sorace’s work by proposing a classification of peripheral verbs. Considering that telicity correlates with ‘be’ and agentivity with ‘have’ in Sorace’s ASH (2000), our proposal establishes a continuum between ‘have’ and ‘be’ (Figure 2).

<b>PERIPHERAL VERBS</b>	[ -AGENTIVITY], [ -INTERNAL CAUSE]	[ $\pm$ HUM]	<b>BE</b>
	[ -AGENTIVITY], [ +INTERNAL CAUSE]		<b>BE/HAVE</b>
	[ +AGENTIVITY], [ +INTERNAL CAUSE]		<b>HAVE</b>

**FIGURE 2** – Proposal for a gradient conceptualization of the agentivity feature as a parameter favoring the auxiliary ‘have’.

In the continuum represented by the peripheral verbs, there are different possibilities. When the subject is nor agentive neither an internal cause, hence properly non-agentive, the auxiliary chosen is ‘be’. When the subject is agentive (and therefore an internal cause), the selection is ‘have’. Finally, when the subject is non-agentive but is an internal cause (thus having both patient and agent characteristics), there is variation. This ambiguity allows the selection of both auxiliaries. ‘Have’ is normally preferred (especially when the feature of internal cause is combined with the [ $\pm$ HUMAN] trait), but the crucial point is that both auxiliaries become possible. This result has been particularly important.

The hierarchy proposed in the present study in terms of binary features especially highlighted the gradience with respect to agentivity: we have indeed [ $\pm$ AGENTIVITY][ $\pm$ INTERNAL CAUSE], [ $\pm$ AGENTIVITY][ $\pm$ INTERNAL CAUSE], and [ $\pm$ AGENTIVITY][ $\pm$ INTERNAL CAUSE], corresponding to the fully intended parameter of agentivity. This model is valid for Italian peripheral verbs, which have been shown to be particularly influenced by semantic parameters. The advantage

of this hierarchy is that it incorporates the concept of gradience while maintaining a symmetric structure with [-AGENTIVITY] on one hand and [+AGENTIVITY] on the other. The configuration [-AGENTIVE][+INTERNAL CAUSE] allows for both ‘have’ and ‘be’. The feature [±HUMAN] should be considered in relation to this configuration: if the subject is [+HUMAN], the likelihood of selecting ‘have’ is higher, and conversely, if the subject is [-HUMAN], ‘be’ becomes more likely.

As to French, although the identified semantic factors are not applicable to all peripheral verbs, they are relevant for one of the sub-classes of these verbs, namely the change-of-state verbs, accounting for the appearance of ‘have’ where ‘be’ would be expected.

Other French verbs, as shown during the analysis, are particularly influenced by construction, demonstrating different behaviors depending on the construction in which they appear (e.g., spatial motion verbs present ‘be’ in spatial motion constructions and ‘have’ in quantitative ones).

## **DISCUSSION OF THEORETICAL IMPLICATIONS**

This thesis contributes significantly to the existing literature on auxiliary selection. It supports Sorace’s notion of a gradient approach to auxiliary selection rather than a strict binary classification of verbs. It even refines this approach by replacing the binary opposition [±AGENTIVE] with a more gradual distinction, incorporating the features [±INTERNAL CAUSE] and [±HUMAN], identified through a corpus study of Italian peripheral verbs. The distinction between internal cause and agentivity, with agentivity involving internal cause and volition, is crucial. Previous studies not focusing on auxiliary selection have independently demonstrated the necessity of identifying agentivity as causation with volition (DeLancey, 1984).

This thesis challenges the universal nature of the ASH, by revealing a clear distinction as to the principles governing auxiliary selection in Italian and French. While in Italian auxiliary selection is conditioned by the semantic features of the argument in the subject position, in French it is conditioned by the construction in which verbs are embedded. This leads to similar patterns for different verbs used in the same construction and different patterns for a single verb when embedded in different constructions, regardless of the semantic features of the subject. This finding is crucial, as it contradicts earlier studies that analyzed French verbs through semantic and aspectual features relevant for Italian (Sorace, 2000; Legendre & Sorace, 2003).

Furthermore, the study’s methodology, combining qualitative and quantitative analyses, provides a model for future linguistic research. It demonstrates the value of an integrative

approach to understanding complex linguistic phenomena and ensures the richness and reliability of the analyses.

## **PERSPECTIVES FOR FUTURE RESEARCH**

Building on the findings of this study, future research could explore auxiliary selection in other languages (Washio 2004; Legendre 2007; Gillmann 2015), to see how these patterns manifest in different linguistic contexts. Additionally, investigating the role of auxiliary selection in spoken language and across different registers would also provide valuable insights.

Further studies could also examine the cognitive processes underlying auxiliary selection, exploring how native speakers and language learners perceive and apply these rules.

Another possibility is to add empirical ground to the confirmed hypotheses by asking to randomly selected native speakers of Italian and French how much they rate a list of sentences containing the different possible combinations of semantic and syntactic features.

Questionnaires represent a complimentary way of further confirming our hypotheses: the focus is on perception (acceptability judgements) instead on language production (written sentences taken from Sketch Engine). This step has already been initiated: questionnaires have been created for Italian verbs, whether functioning as full lexical verbs or as semi-auxiliaries, and for French verbs through a factorial design encompassing the most important factors. For Italian verbs, three representative verbs were chosen (*cambiare* for predominantly ‘be’ selection verbs, *contare* for predominantly ‘have’ selection verbs and *suonare* for verbs allowing both auxiliaries). These verbs were combined with parameters that have proven relevant through corpus and statistical analyses, including semantic traits for full verbs and combinations of the different infinitive types with semantic traits for semi-auxiliaries. For French verbs, one verb from each class was chosen: *monter* (spatial motion), *changer* (change of state), and *apparaître* (appearance verbs).

Based on this, two questionnaires were designed for each category (Italian full verbs, Italian semi-auxiliaries, and French full verbs) using a Latin Square design (Bross, 2019). The sentences were subjected to acceptability judgments. Although the results could not be included in this thesis, the questionnaires for Italian verbs, which involved 392 and 382 participants, confirmed the analyses conducted so far, and further corroborated the proposed hypotheses.

## **FINAL REMARKS**

This thesis provides a detailed and nuanced investigation of auxiliary selection in Italian and French, highlighting the differences and similarities between the two languages. Crucial to this study is the relevance of the concept of internal cause for Italian verbs and the importance of construction types for French verbs.

The integrative methodological approach has proven effective in uncovering the complexities of this linguistic phenomenon, demonstrating that combining innovative methods with theoretical analysis enriches the explanation. The hypotheses derived from the manual annotation of sentences and quantitative statistical methods confirm the completeness and reliability of this analysis.

The findings contribute to a deeper understanding of verb categorization and auxiliary choice, with significant implications for both theoretical linguistics and practical applications. As we continue to delve into the rich tapestry of language, studies like this remind us of the intricate connections that shape how we communicate and understand the world around us.



## REFERENCES

- Ackema, P., & Sorace, A. (2017). Auxiliary Selection. In M. Everaert & H. C. van Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax, Second Edition* (pp. 1–32). Hoboken, New Jersey: John Wiley & Sons, Inc.  
<https://doi.org/10.1002/9781118358733.wbsyncom072>
- Alexiadou, A., Anagnostopoulou, E., & Everaert, M. (2004). *The Unaccusativity Puzzle*. Oxford: Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780199257652.001.0001>
- Amato, I. (2022). Auxiliary selection is Agree: Person-driven and argument-structure-based splits. *Isogloss. Open Journal of Romance Linguistics*, 8(2), 1–20.  
<https://doi.org/10.5565/rev/isogloss.131>
- Aranovich, R. (2003). The grammar of the Spanish perfect tense. *Studies in Linguistics* 3(2), 123-145.
- Arunachalam, S. (2013). Experimental Methods for Linguists: Experimental Methods for Linguists. *Language and Linguistics Compass* 7(4), 221–232.  
<https://doi.org/10.1111/lnc3.12021>
- Barbiers, S., & Sybesma, R. (2004). On the different verbal behavior of auxiliaries. *Lingua*, 114(4), 389–398. [https://doi.org/10.1016/S0024-3841\(03\)00065-2](https://doi.org/10.1016/S0024-3841(03)00065-2)
- Bentley, D., & Eythórsson, T. (2004). Auxiliary selection and the semantics of unaccusativity. *Lingua*, 114(4), 447–471. [https://doi.org/10.1016/S0024-3841\(03\)00068-8](https://doi.org/10.1016/S0024-3841(03)00068-8)
- Blumenthal, P. (1990). Classement des adverbes: Pas la Couleur, rien que la nuance? *Langue française* 88(1), 41–50. <https://doi.org/10.3406/lfr.1990.5751>
- Breiman, L. (2001). Random forests. *Machine Learning* 45(1), 5-32.
- Bross, F. (2019). The Latin Square design in linguistic research. *Journal of Experimental Linguistics* 56(1), 112-130.
- Buchard, A., & Carlier, A. (2008). La forme verbale « être + participe passé » en tant que marqueur d'aspect et de structure argumentale : une typologie graduée. In *Actes du 1er Congrès Mondial de Linguistique Française* (pp. 210-222). Paris: EDP Sciences.
- Burzio, L. (1986). *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel.

- Burzio, L. (1981). *Intransitive Verbs and Italian Auxiliaries*. Cambridge, Mass.: MIT Press.
- Carlier, A., & Sarda, L. (2010). Le complément de la localisation spatiale: Entre argument et adjectif. *2ème Congrès Mondial de Linguistique Française*, 138. <https://doi.org/10.1051/cmlf/2010251>
- Cinque, G. (2004). “Restructuring” and functional structure. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures* (Vol. 3, pp. 132-191). Oxford University Press.
- Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press.
- Crisma, P., & Longobardi, G. (2009). *Historical Syntax and Linguistic Theory*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199560547.001.0001>
- Cruse, D. A. (1973). Some Thoughts on Agentivity. *Journal of Linguistics*, 9(1), 11–23. JSTOR.
- D’Alessandro, R. (2017). When you have too many features: Auxiliaries, agreement and clitics in Italian varieties. *Glossa: a journal of general linguistics*, 2(1). <https://doi.org/10.5334/gjgl.102>
- DeLancey, S. (1984). Notes on Agentivity and Causation. *Studies in Language*, 8(2), 181–213. <https://doi.org/10.1075/sl.8.2.05del>
- Demirdache, H., & Martin, F. (2015). Agent control over non culminating events. In E. Barrañón López, J. L. Cifuentes Honrubia, & S. Rodríguez Rosique (Eds.), *IVITRA Research in Linguistics and Literature* (Vol. 9, pp. 185–217). John Benjamins. <https://doi.org/10.1075/ivitra.9.09dem>
- Dowty, D. (1991). Thematic Proto-Roles and Argument Selection. *Language*, 67(3), 547. <https://doi.org/10.2307/415037>
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague’s PTQ*. D. Reidel Pub. Co.
- Dufter, A., & Stark, E. (Eds.). (2017). *Manual of Romance morphosyntax and syntax*. De Gruyter.
- Durand, J., Habert, B., & Laks, B. (Eds.). (2008). *Congrès mondial de linguistique française: Paris, 9-12 juillet 2008 : recueil des résumés* [CD-ROM des actes]. EDP Sciences.

Fellbaum, C., & Pustejovsky, J. (1997). The Generative Lexicon. *Language*, 73(3), 597. <https://doi.org/10.2307/415891>

Finocchiaro, C. (2002). Sensitivity to the verb [ $\pm$ agentive] feature: The case of an aphasic subject. *Journal of Neurolinguistics*, 15(3–5), 433–446. [https://doi.org/10.1016/S0911-6044\(01\)00033-1](https://doi.org/10.1016/S0911-6044(01)00033-1)

Flaux, N., & Van de Velde, D. (2000). *Les noms en français, esquisse de classement*. Ophrys.

Fritz-Huechante, P., Verhoeven, E., & Rott, J. A. (2020). Agentivity and non-culminating causation in the psych domain: Cross-linguistic evidence from Spanish and Korean. *Glossa: a journal of general linguistics*, 5(1). <https://doi.org/10.5334/gjgl.896>

Giancarli, P.-D. (2015). Auxiliary selection with intransitive and reflexive verbs: The limits of gradience and scalarity, followed by a proposal. In R. Kailuweit & M. Rosemeyer (Eds.), *Auxiliary Selection Revisited* (pp. 79–120). De Gruyter. <https://doi.org/10.1515/9783110348866-004>

Gillmann, M. (2015). Auxiliary selection in closely related languages: The case of German and Dutch. In R. Kailuweit & M. Rosemeyer (Eds.), *Auxiliary Selection Revisited* (pp. 333–358). De Gruyter. <https://doi.org/10.1515/9783110348866-012>

Godard, D., & Abeillé, A. (Eds.). (2003). *Les langues romanes: Problèmes de la phrase simple*. CNRS.

Gregersen, F., Engdahl, E., & Laanemets, A. (2017). Introduction to the special issue on variation in auxiliary selection. *Acta Linguistica Hafniensia*, 49(2), 107–117. <https://doi.org/10.1080/03740463.2017.1380252>

Grimshaw, J. (1990). *Argument Structure*. MIT Press.

Heidinger, S. (2015). The auxiliary selection of French *monter* ‘move upward’ from the 16<sup>th</sup> to the 20<sup>th</sup> century. In R. Kailuweit & M. Rosemeyer (Eds.), *Auxiliary Selection Revisited* (pp. 277–300). De Gruyter. <https://doi.org/10.1515/9783110348866-010>

Heine, B. (1993). *Auxiliaries: Cognitive forces and grammaticalization*. Oxford University Press.

Heine, B., & Kuteva, T. (2006). *The Changing Languages of Europe*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199297337.001.0001>

Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen corpus family. *7th International Corpus Linguistics Conference*, 125–127.

Kailuweit, R., & Rosemeyer, M. (Eds.). (2015). *Auxiliary selection revisited: Gradience and gradualness*. De Gruyter.

Karssenber, L. (2018). Variation in French impersonal constructions. *Language Variation and Change*, 30(2), 235-259.

Kass, G. V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, 29(2), 119. <https://doi.org/10.2307/2986296>

Kayne, R. S. (1993). Toward a modular theory of auxiliary selection. *Studia Linguistica*, 47(1), 3-31.

Keller, F., & Sorace, A. (2003). Gradient auxiliary selection and impersonal passivization in German: An experimental investigation. *Journal of Linguistics*, 39(1), 57–108. <https://doi.org/10.1017/S0022226702001676>

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1-26.

Ledgeway, A. (2019). Parameters in the development of Romance perfective auxiliary selection. In M. Cennamo & C. Fabrizio (Eds.), *Current Issues in Linguistic Theory* (Vol. 348, pp. 344–384). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.348.17led>

Legendre, G. & Sorace, A. (2003). Auxiliaries and intransitivity in French and in Romance. In D. Godard (Ed.), *Les langues romanes : Problèmes de la phrase simple* (English version of ‘Auxiliaires et intransitivité en français et dans les langues romanes’). Paris: CNRS Edition.

Legendre, G. (2007). On the typology of auxiliary selection. *Lingua* 17 (2007), 1522–1540.

Levin, B., & Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge University Press.

Levin, B., & Rappaport Hovav, M. (1995). *Unaccusativity: At the Syntax-Lexical Semantics Interface*. MIT Press.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press.

Liaw, A., & Wiener, M. (2002). Classification and regression by RandomForest. *R News*, 2(3), 18-22.

Lieber, R., & Baayen, H. (1997). A Semantic Principle of Auxiliary Selection in Dutch. *Natural Language & Linguistic Theory*, 15(4), 789–845.

Loporcaro, M. (2003). The Unaccusative Hypothesis and participial absolutes in Italian. Perlmutter's generalization revised. <https://doi.org/10.5167/UZH-221877>

Lorusso, P. (2015). Auxiliaries and Verb Classes in Child Italian: A Syntactic Analysis of the Development of Aspect. *Quaderni Di Linguistica e Studi Orientali*, 61-88 Pages. <https://doi.org/10.13128/QULSO-2421-7220-16516>

Mačutek, J., & Wimmer, G. (2013). Evaluating goodness-of-fit of discrete distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, 20(3), 227–240. <https://doi.org/10.1080/09296174.2013.799912>

Martin, F. (2020). Aspectual Differences Between Agentive and Non-agentive Uses of Causative Predicates. In E. A. Bar-Asher Siegal & N. Boneh (Eds.), *Perspectives on Causation* (pp. 257–294). Springer International Publishing. [https://doi.org/10.1007/978-3-030-34308-8\\_8](https://doi.org/10.1007/978-3-030-34308-8_8)

Martin, F., & Schäfer, F. (2017). *Sublexical modality in defeasible causative verbs (Vol. 1)*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198718208.003.0006>

Mateu, J. (2009). Split intransitivity in Catalan. *Catalan Journal of Linguistics*, 8(1), 35-58.

McFadden, T. (2007). Auxiliary Selection. *Language and Linguistics Compass*, 1(6), 674–708. <https://doi.org/10.1111/j.1749-818X.2007.00034.x>

McLendon, S. (1978). Ergativity, Case, and Transitivity in Eastern Pomo. *International Journal of American Linguistics*, 44(1), 1–9. <https://doi.org/10.1086/465512>

Nelson, D. (2000). Miriam Butt & Wilhelm Geuder (eds.), The projection of arguments: Lexical and compositional factors (CSLI Lecture Notes 83 ). Stanford: CSLI Publications, 1998. Pp. viii+363. *Journal of Linguistics*, 36(1), 157–199. <https://doi.org/10.1017/S0022226799228083>

Perlmutter, D. M. (1989). Multiattachment and the Unaccusative Hypothesis: The Perfect Auxiliary in Italian. *Probus*, 1(1), 63-119.

Perlmutter, D. M. (1979). Universal Grammar and the Syntax of Spanish Reflexives. In T. Givón (Ed.), *Syntax and Semantics* (Vol. 12, pp. 171-213). Academic Press.

Perlmutter, D. M. (1978). Impersonal Passives and the Unaccusative Hypothesis. *Annual Meeting of the Berkeley Linguistics Society*, 4, 157. <https://doi.org/10.3765/bls.v4i0.2198>

Pountain, C. (1985). Copulas, verbs of possession and auxiliaries in Old Spanish: the evidence for structurally interdependent changes. *Bulletin of Hispanic Studies* 62 (4), 337-355. <https://doi.org/10.1080/1475382852000362337>

- Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press.
- Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The Projection of Arguments: Lexical and Compositional Factors* (pp. 97-134). Stanford, CA: CSLI Publications.
- Rea, B. (2018). Le retour d'un « être » cher ? Étude sociolinguistique de l'alternance des auxiliaires dans le français de Montréal (1971-2016). *SHS Web of Conferences*, 46, 02002. <https://doi.org/10.1051/shsconf/20184602002>
- Reinhart, T. (2003). The Theta System – an overview. *Theoretical Linguistics*, 28(3). <https://doi.org/10.1515/thli.28.3.229>
- Reinhart, T. (2016). The Theta System: Syntactic Realization of Verbal Concepts. In M. Everaert & M. Marelj (Eds.), *Concepts, Syntax, and Their Interface* (pp. 1–112). The MIT Press. <https://doi.org/10.7551/mitpress/9780262034135.003.0001>
- Rizzi, L. (1982). *Issues in Italian syntax*. Studies in Generative Grammar, 11. Foris.
- Rizzi, L. (1976). Ristrutturazione. *Rivista di grammatica generativa*, 1, 1-54.
- Ritschard, G. (2013). CHAID and Earlier Supervised Tree Methods. In J. J. McArdle & G. Ritschard (Eds.), *Contemporary Issues in Exploratory Data Mining in Behavioral Sciences* (pp. 48–74). Routledge, New York.
- Rosemeyer, Malte. (2014). *Auxiliary Selection in Spanish. Gradience, Gradualness, and Conservation*. Amsterdam, Philadelphia: John Benjamins.
- Rosen, C. (1981). The Relational Structure of Reflexive Clauses: Evidence from Italian. *Linguistic Inquiry*, 12(1), 85-95.
- Schlesinger, I. M. ([1995] 2006). *Cognitive Space and Linguistic Case: Semantic and Syntactic Categories in English*. Cambridge: Cambridge University Press.
- Smessaert, H., & ter Meulen, A. G. B. (2004). Temporal Reasoning with Aspectual Adverbs. *Linguistics and Philosophy*, 27(2), 209–261.
- Sorace, A. (2015). The cognitive complexity of auxiliary selection: From processing to grammaticality judgements. In R. Kailuweit & M. Rosemeyer (Eds.), *Auxiliary Selection Revisited* (pp. 23–42). De Gruyter. <https://doi.org/10.1515/9783110348866-002>
- Sorace, A. (2000). Gradients in Auxiliary Selection with Intransitive Verbs. *Language*, 76(4), 859. <https://doi.org/10.2307/417202>

Talmy, L. (1988). Force Dynamics in Language and Cognition. *Cognitive Science*, 12(1), 49–100. [https://doi.org/10.1207/s15516709cog1201\\_2](https://doi.org/10.1207/s15516709cog1201_2)

Tenny, C. L. (1994). *Aspectual Roles and the Syntax-Semantics Interface* (Vol. 52). Springer Netherlands. <https://doi.org/10.1007/978-94-011-1150-8>

Torcolacci, G. (2015). Marking the default: *Auxiliary selection in Southern Italian dialects* (LOT No. 385). *LOT*.

ter Meulen, A. G. B. (2004). The dynamic semantics of aspectual adverbs. *Empirical Issues in Formal Syntax and Semantics* 5, 241–253.

Van de Velde, D. (1995). Le spectre nominal : *Des noms de matières aux noms d'abstractions*. Peeters.

Van Valin, R. D. (1990). Semantic Parameters of Split Intransitivity. *Language*, 66(2), 221. <https://doi.org/10.2307/414886>

Van Valin, R. D. (1993). *Advances in role and reference grammar*. J. Benjamins Pub. Co.

Van Voorst, J. (1992). The Aspectual Semantics of Psychological Verbs. *Linguistics and Philosophy*, 15(1), 65–92. JSTOR.

Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, 66(2), 143. <https://doi.org/10.2307/2182371>

Vikner, S. (1989). Topics in Scandinavian Syntax. *Nordic Journal of Linguistics*, 12(1), 79–90. <https://doi.org/10.1017/S0332586500001943>

Vlach, F. (1993). Temporal adverbials, tenses and the perfect. *Linguistics and Philosophy*, 16(3), 231–283. <https://doi.org/10.1007/BF00985970>

Washio, R. (2004). Auxiliary Selection in the East. *Journal of East Asian Linguistics*, 13(3), 197–256. <https://doi.org/10.1023/B:JEAL.0000038249.86375.a5>

Williams, E. (1997). Beth Levin & Malka Rappaport Hovav, Unaccusativity: At the syntax-lexical interface. Cambridge, MA: MIT Press, 1995. Pp. xii+336. *Journal of Linguistics*, 33(2), 563–641. <https://doi.org/10.1017/S0022226796276571>

Zipf, G. K. (1949). Human behavior and the principle of least effort. *Journal of Consulting Psychology*, 13(3), 224–224. <https://doi.org/10.1037/h0052803>