

# *French ingressives and (phasal) aspect: A frame-semantic corpus-based analysis*

FILIP VERROENS 

Ghent University, Flanders, Belgium

[Filip.Verroens@UGent.be](mailto:Filip.Verroens@UGent.be)

and

LUDOVIC DE CUYPERE

Ghent University, Flanders, Belgium

Vrije Universiteit Brussel, BCLS, Flanders, Belgium

[Ludovic.DeCuyper@UGent.be](mailto:Ludovic.DeCuyper@UGent.be)

---

## **Abstract**

This article compares the usage of *commencer à* ‘to begin’+Vinf. and *se mettre à* ‘to start’ + Vinf. in modern French. Using a corpus sample of 2000 observations, we examined the effect of Adverbial complementation, of Event type (aspect), and of Tense. Based on a mixed-effects logistic regression analysis, we found evidence for Event type, where *se mettre à* is associated with activities, and Tense, where *se mettre à* seems to be associated with passé simple, futur proche and subjonctif présent. On the other hand, *commencer à* is associated with plus-que-parfait and indicatif imparfait. We discuss the results in the frame-semantic model of Croft (2012). We make the case that *commencer à* can have the profile of an achievement or that of an accomplishment, while *se mettre à* manifests only one profile, that of an achievement. Our results support a one-component approach to aspect in which the result of the interaction between grammatical aspect and lexical aspect can be attributed to the same aspectual contour.

**Keywords:** French ingressives, aspect, mixed-effects logistic regression analysis, frame-semantics

## **Résumé**

Cet article compare l’usage de *commencer à* +Vinf. et *se mettre à* + Vinf. en français moderne. À l’aide d’un corpus de 2000 occurrences, nous avons examiné l’effet de la complémentation adverbiale, du type de procès et du temps verbal chez les deux constructions. Ayant utilisé la méthode de la régression logistique à effets mixtes, nous avons trouvé des résultats significatifs pour le Type de procès – *se mettre à* est associé à des activités – et pour le Temps verbal – *se mettre à* est associé au passé simple, au futur proche et au subjonctif présent, tandis que *commencer à* est associé au plus-que-parfait et à l’indicatif imparfait. Nous discutons les résultats dans le modèle sémantique de Croft (2012) et avançons que *commencer à* peut avoir le profil d’un achèvement ou celui d’un accomplissement alors que *se mettre à* manifeste un seul profil,

à savoir celui d'un achèvement. Nos résultats soutiennent une approche unidimensionnelle de l'aspect dans laquelle le résultat de l'interaction entre l'aspect grammatical et l'aspect lexical peut être attribué au même contour aspectuel.

**Mots clés:** inchoatifs français, aspect, régression logistique à effets mixtes, sémantique des cadres

## 1. INTRODUCTION

The French ingressive<sup>1</sup> constructions *commencer à* (COMA) 'to begin' + *Vinf.* and *se mettre à* (SMA) 'to start' + *Vinf.*<sup>2</sup> are commonly considered to be synonyms in reference grammars (e.g., Wilmet 1998: 321) and are, at first sight, used interchangeably<sup>3</sup>:

- (1) Bientôt, l'ETA commence à regarder du côté du tiers-monde, se met à parler de  
« guerre révolutionnaire ». (LM111)  
'Soon, the ETA begins to look towards the Third World, starts to speak of "revolutionary war".'

Nevertheless, several researchers have noted differences in their usage patterns. A first extra-linguistic difference is related to register: SMA appears to be rarer in journalistic texts (Roy 1976: 284, Peeters 1993: 41–42). Previous scholarship<sup>4</sup> has additionally noted a possible relation with three linguistic factors: event type (2), adverbial complementation (3)–(4), and tense (5).

- (2) Le chien de nos voisins [commença/ \*se mit] à être sourd. (Peeters 1993: 40)  
'Our neighbor's dog [began / \*started] to be deaf.'
- (3) Le soldat amnésique [se met soudain à/ ?commence soudain à] chanter. (Sato 1994: 31)  
'The amnesic soldier [suddenly starts /? suddenly begins] to sing.'
- (4) Je [n'ai pas commencé/ ? ne me suis pas mis] à manger. (Franckel 1989: 144)  
'I [haven't begun /? haven't started] to eat.'
- (5) Je [commencerais / ?? me mettrais] à travailler. (Sato 1994: 32).  
'I will [begin / ?? start] to work.'

Several authors have noticed that a stative constraint (2) appears when aspectual constructions combine with states (e.g., Lamiroy 1987, Peeters 1993, Marque-Pucheu 1999). Lamiroy (1987) proposes several syntactic conditions likely to cancel the stative constraint: the use of plural, collective or generic subject NPs, infinitival

<sup>1</sup>On the notions of ingressive/inchoative in French linguistics, see Verroens (2018).

<sup>2</sup>Throughout the article, we have translated SMA as 'to start' and COMA as 'to begin'. We are aware that the inter- and intra-linguistic differences are not straightforward, but we have opted for this consistent translation for the sake of simplicity.

<sup>3</sup>Abbreviations used: COMA: *commencer à...*; FT: Frantext corpus; LM: *Le Monde* corpus; MDS: multidimensional scaling; SMA: *se mettre à...*

<sup>4</sup>In particular, we refer to enunciative analyses (Franckel 1989, Sato 1994, Saunier 1999), analyses in Natural Semantic Metalanguage (Peeters 1993) and logical analyses (Nef 1980, Gardiès 1981, Marque-Pucheu 1999).

complements with *se-moyen* (“middle” reflexive construction), the use of adverbial complements and of the complementizer *par*. What characterizes these different contexts is that they no longer refer to a single state, but to a state that is repeated, usual or linked to another event. Lamiroy’s hypothesis (1987: 293) states that the stative constraint will be less pronounced for sequentially interpreted expressions than for aspectual verbs. Peeters (1993) claims that the stative constraint for SMA disappears in certain contexts, more precisely with an iterative reading, with another verb intervening, or if there is a clear action element in glosses with a state predicate. For SMA in particular, it is generally accepted that it differs from COMA in that it cannot be combined with this event type (2). With regard to the combination with achievements, several studies (Verbert 1979, Peeters 1993, Iordache and Scurtu 1994) have shown that it would be difficult, if not impossible, given that it is a punctual change which prevents SMA and COMA from marking the onset, as they normally do with durative events. As for the stative constraint several (more or less similar) phenomena are likely to facilitate the combination with achievements, generating a perspective of “slow-motion” (Verbert 1979: 71), a sequential interpretation (Lamiroy 1987: 293), or the effect of “imperfectivising perfectives” (Wilmet 1998: 317–318): with a plural or collective subject NP, a temporal subordinate clause or with a metaphorical use of the achievement predicate.

With regard to the second constraint, it has been noticed that SMA is usually associated with adverbs of velocity and suddenness (e.g., *soudain*, *tout à coup*, *brusquement*, etc. ‘suddenly, abruptly’) (Coseriu 1976, Peeters 1993, Saunier 1999) and that this construction marks a more “brutal” inception than COMA which generates a more “attenuated” inceptive value (Franckel 1989: 147). Sato (1994) goes so far as to suggest that the distribution of this type of adverbs is exclusively reserved for SMA, in other words that they cannot appear with COMA, as shown in (3). Peeters (1993) and Saunier (1999) observe however that the use of adverbs of suddenness is not systematic and that opposite evidence exists. As for the negation expressed in (4), Franckel (1989: 144) judges that the negation would be very strongly constrained, if not even impossible for SMA. While the negation constraint is rejected by Peeters (1993: 43), Franckel (1989) and Saunier (1999) explain the absence of negation by the aoristic functioning of SMA.<sup>5</sup>

Finally, Nef (1980) and Sato (1994) report a future constraint (5) for SMA, but not for COMA. According to the latter author, it is very difficult, if not impossible, to have SMA in the future for the simple reason that it would violate the unexpectedness of SMA, while COMA implies anticipation or intentionality. In general, the verbal tense of inceptive constructions has aroused little interest and is limited to a few observations. We will look at the distribution of all moods and tenses and examine whether there really is a constraint of the future for SMA as suggested by Sato (1994), or if this constraint occurs rather with COMA (Nef 1980), or if COMA

<sup>5</sup>Saunier (1999: 279) paraphrases Franckel by saying that *do not P* supposes that *P* has a status in relation to which the predication is inscribed. Insofar as SMA introduces the relation of a subject to an event as having no status apart from his temporal localization, the negation cannot relate to this localization, but only to the veracity of the predication taken as a whole.

clearly privileges the imperfect (Nef 1980). These supposed temporal restrictions are interesting because they are often used as an argument to distinguish lexical periphrases, to which our ingressives belong, from grammatical periphrases which are not compatible with all tenses (Laca 2004: 90, Laca 2005: 51, Vet 2008: 458).

To date, little empirical research exists that evaluates to what extent the SMA vs. COMA alternation is associated with the mentioned constraints. Indeed, as illustrated by the counterexamples in (6)–(9), opposite to (2)–(5) respectively, the semantic contrast is not categorical.

(6) Le café de Barcelone se mettait à ressembler à une espèce de hall de gare. (FT009)

‘The Barcelona bar started to look like a kind of railway station.’

(7) Jeff, tout à coup, commence à chanter « Malbrough s’en va-t-en guerre ». (FT060)

‘Jeff suddenly begins to sing “Malbrough s’en va-t-en guerre”.’

(8) Aucune confédération ne s’est mise à revendiquer une sixième semaine, préférant donner la priorité aux 35 heures. (LM080)

‘No confederation has started to claim a sixth week, preferring to give priority to the 35 hours.’

(9) Les cloches de la basilique se mettront au même moment à sonner à toute volée. (FT106)

‘At the same moment, the bells of the basilica will start to ring out loud and clear.’

The present study offers a corpus-based analysis. We additionally examine the consequences of our results for phasal aspect and the concept of aspect in relation to the frame-semantic model of Croft (2012) which presents an integrated analysis of aspect and force-dynamics in a single representation of verbal semantics. By *phase* we understand a temporal part of an event defined as unfolding over time, which has a single aspectual construal for that time period (Croft 2012: 403). More precisely, we will be concerned with the initial transition phase of an event. With regard to the notion of aspect, there is still a lot of discussion. De Swart (2012: 766) notes that

Some researchers defend a two-component theory in which grammatical aspect and aspectual class are interpreted by means of different sets of tools (Smith 1997; Depraetere 1995; Filip 1999; Bertinetto and Delfitto 2000). Others use the same semantic machinery for both, whichever that may be (Moens and Steedman 1988, Parsons 1990, Kamp and Reyle 1993, de Swart 1998, Verkuyl 1999, Cipria and Roberts 2000). What both lines of analysis agree on is that grammatical aspect determines the aspectual nature of the sentence as a whole, and may overrule certain semantic features of its internal aspectual make-up.

Following Croft (2009, 2012), we adopt a one-component approach stipulating that the semantics of grammatical aspect is the same as that of lexical aspect: grammatical aspect interacts with lexical aspect, but the result is of the same semantic type as lexical aspect. In two-component approaches, on the other hand, grammatical aspect is considered to be semantically distinct from lexical aspect. Its semantic structure is accordingly considered as different and it is generally characterized as a



**Figure 1:** Alternative profiling of English *see* (Croft 2012: 55)

“viewpoint” semantics (Sasse 2002: 202–203, Michaelis 2004: 9, Croft 2012: 31–32). In the frame-semantic representation of Croft (2012: 31–69), events are represented in two dimensions (Figure 1): t-dimension (temporal boundaries) and q-dimension (qualitative/inherent boundaries referring to the presence of a result state/telos). Punctual events are t points, while durative events extend on t. Stative events are points on q, while dynamic events extend on q (representing changes from one qualitative state to another). Predicates can be given more than one interpretation (i.e., they can be presented through more than one construal). Thus, verbs of cognition (e.g., *to understand*) and of physical perception (e.g., *to see*) can be analyzed as states or achievements (Vendler 1967: 113–119). For example, the frame-semantic representation of the verb *to see* (Figure 1a–b) changes depending on the profiled phase of the event. In (10a), it is the resulting seeing state which is profiled by the verb + Simple Present construction; in (10b), it is the achievement that is profiled through the verb + Simple Past.

- (10) a. I see Mount Tamalpais. (Croft 2012: 54)  
 b. I reached the crest of the hill and saw Mount Tamalpais. (Croft 2012: 54)

The analysis of the predicate into a particular event type is therefore not given, but depends on the choice of the speaker among the different ways of conceptualizing and describing a scene (i.e., construal). With this set of basic distinctions, as well as with the cognitive notions of profile and aspectual contour, that is, the sequence of phases representing how a particular event is construed as unfolding over time (Croft 2012: 398), we can represent the semantics of the French ingressives. A verb in a particular grammatical context designates or profiles one (or more) phases of the aspectual contour of an event. In this model of representation, both the profile and the aspectual contour are part of the meaning of linguistic form. Representations in t / q phases allow for the incorporation of interpretative and aspectual distinctions identified in aspectual literature since Vendler (1967). Croft’s classification (Croft 2012: 31–69) is an elaboration of the four-way Vendler classification and can be resumed as follows:

- **States:** (a) inherent permanent states (*She is French*), (b) acquired permanent states (*The window is shattered*), (c) transitory states (*The door is open*), (d) point states (*The sun is at its zenith*)
- **Activities:** (a) directed activities (*The soup cooled*), (b) undirected (cyclic) activities (*The girls chanted*)
- **Accomplishments:** (a) incremental accomplishments (*I ate an apple cake*), (b) non-incremental accomplishments (runup achievements) (*Harry repaired the computer*)

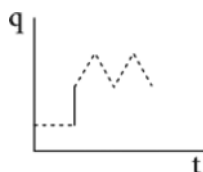
- **Achievements:** (a) (directed) reversible achievements (*The door opened*), (b) (directed) irreversible achievements (*The window shattered*), (c) (undirected) cyclic achievements (semelfactives) (*The mouse squeaked*).

For the quantitative analysis (section 3), we have used the classic Vendler classification, as Croft's finer-grained classification is likely to lead to categories too small to infer statistically significant patterns. Nevertheless, some finer-grained distinctions (e.g., permanent vs. transitory states) will be relevant for the qualitative analysis (section 4).

Ingressive constructions alter the aspectual structure of these basic events. As de Swart (1998), Laca (2002, 2004) and Michaelis (2004) note, constructions that modify the aspectual structure of an event also generally require that this event be interpreted as a particular type of aspect, which is then modified when combined with the construction modifying this aspect: compare 'He ate the apples' [accomplishment] with 'He started eating the apples' [achievement] (Laca 2002: 72). French aspectual constructions include auxiliary verbs, aspectual adverbs, and basic tense constructions. In this article, we examine the interaction of the two ingressive constructions with aspectual interpretations of predicates, the semantic analysis of adverbial constructions and tenses. Croft's framework provides a natural representation for a variety of constructions that alter the aspectual structure of events. With regard to English ingressives, Croft (2012: 106) acknowledges that subtle semantic differences exist between *to start* and *to begin*, but he doesn't attempt to identify them. He remarks that inherent permanent states are generally not construable with ingressive constructions, since they lack transitions (*\*She started/began to be Ukrainian*) and analyzes *She started to dance* as a directed achievement (Figure 2) in which the inception transition phase is profiled by the ingressive construction and followed by a basic event (state or process) as indicated by the unprofiled activity following the transition phase.

With respect to French ingressives, we expect clear aspectual differences (in terms of a punctual/durative analysis) between COMA and SMA constructions with regard to all the variables under study; we hypothesize that inherent permanent states are not impossible with ingressive constructions, and suppose that these two ingressives will profile the inceptive transition phase in a different way.

As for the status of our ingressive constructions, several views exist in the literature. Some researchers (Imbs 1960, Martin 1971, Wagner and Pinchon 1973), relate aspectual periphrases to lexical aspect, while others (Gross 1975, Gaatone 1995, Barceló and Bres 2006), relate them to grammatical aspect. For still others (Lamiroy 1999; Laca 2004, 2005; Gosselin 2010a, b, 2011, 2020), some



**Figure 2:** Profiling the inception transition phase (Croft 2012: 106)

constructions relate to lexical aspect and others to grammatical aspect. In this study, we start from the semantic-cognitive analysis of Gosselin that distinguishes between “coV-MAP” (*coverbes de mouvement, de modalité d’action ou de phase*) and “Aux-VA” (*auxiliaires de visée aspectuelle*). The coV-MAPs select the phases of an event by means of categorization. In other words, they construct sub-events which have the value of referential entities, while the Aux-VA mark the aspectual viewpoint, a reference interval which delimits a zone of visibility on the events or sub-events previously constructed. This leads to the following classification (Gosselin 2011: 162):

- **Aux-VA**; *aller, être sur le point/en passe/en train de, (en) être à, venir de, être/avoir + Past Participle.*
- **coV-MAP**; *être prêt à, s’apprêter à, se préparer à, se disposer à, se proposer de, hésiter à, tarder à, partir, monter, s’installer à, aller, s’arrêter à, s’acharner à, s’efforcer de, se dépêcher de, mettre du temps à, être long à, commencer à, se mettre à, continuer de, être occupé à, persévérer à, cesser de, (s’)arrêter de, finir de, (re)venir de, rentrer de, (re)descendre de, sortir de.*

This hypothesis offers a semantic explanation based on syntactic analyses elaborated by other linguists.<sup>6</sup> The considerable variability of the constraints on the different coV-MAPs is generally interpreted as an indication of the different degrees of grammaticalization leading to aux-VA (Squartini 1998, Lamiroy 1999). Ingressives meet the criteria for coV-MAP, although Verroens (2011) argued that they too are partially grammaticalized.<sup>7</sup>

The article is organised as follows. Section 2 outlines the corpus-based methodology. The quantitative results are presented in section 3, while section 4 contains a fine-grained qualitative analysis. Section 5 offers a frame-semantic account for the corpus findings. The conclusions are presented in section 6.

## 2. METHODOLOGY

This section provides information about our corpus and corpus sample, the annotation of our data with respect to the hypotheses explained in section 1, and the statistical model we used to analyse our data sample.

### 2.1 Corpus sample

Our data sample is drawn from the Frantext (FT) literary base for the period 1985 to 2000, and from the journalistic corpus of Le Monde (LM) on CD-ROM (10 / 2004-9 / 2006) from which we have selected the period January 2005 to September 2006. We collected N = 2000 observations: n = 500 occurrences per construction (SMA and

<sup>6</sup>More precisely Kronning (2003), François (2003) and Laca (2005) cited in Gosselin (2011).

<sup>7</sup>Verroens (2011) studies the different parameters of grammaticalization, especially the parameters of scope, semantic integrity, paradigmatic cohesion and paradigmatic variability. Given its distributional and semantic constraints, SMA is found to be less grammaticalized than COMA. See also Willems and Verroens (2019).

COMA) per corpus (LM and FT). Note that the number of COMA is greater than that of SMA in the two corpora. For better comparison, we have balanced the corpus, that is to say we have limited it to the same number of occurrences as SMA. While in Frantext's literary corpus, 4,392,709 words yielded 500 tokens, 34,738,595 were needed in *Le Monde* to obtain the same result. This observation suggests that SMA is indeed rarer in a journalistic corpus.<sup>8</sup>

## 2.2 Data annotation

Every case in our corpus sample was annotated for three variables: Event type, Adverb type, and Mood and Tense.

### 2.2.1 Event type (*state, activity, accomplishment, achievement*)

We classified all first infinitives in one of the four Vendlerian categories (Vendler 1967): *states* (e.g., *savoir* 'to know'), *activities* (e.g., *jouer* 'to play'), *accomplishments* (e.g., *lire un livre* 'to read a book') and *achievements* (e.g., *sortir* 'to leave', *tomber* 'to fall'). Sometimes there were several co-occurring infinitives, but the infinitives 2 and 3 are not part of the actual quantitative analysis as for instance in *Mais, tôt ou tard, il se mettait à consulter ses fiches, chercher, fouiner pour satisfaire le client.* (FT081) 'But, sooner or later, he started consulting his files, searching, nosing round to satisfy the customer.' We expect that the two ingressesives combine with all the event types but that achievements and states will be much less frequent and only possible in specific contexts (see section 1).

### 2.2.2 Adverb type

For each sentence we first annotated the presence of an adverb phrase (present vs. absent). Where an adverb phrase was present, we distinguished between two semantic categories: "velocity" and "suddenness", which we expect to be associated with SMA, and "duration and progression", which we expect to be associated with COMA. Examples of the first category include phrases like *soudain, tout à coup, brusquement*, etc. ('suddenly, abruptly'), examples of the second category are *progressivement, de plus en plus*, etc. ('gradually, more and more').

### 2.2.3 Mood and Tense

We annotated the following Moods and Tenses: conditionnel passé, conditionnel présent, futur antérieur, futur proche, futur simple, impératif, indicatif imparfait, indicatif présent, infinitif, infinitif passé, participe présent, passé antérieur, passé composé, passé récent, passé simple, plus-que-parfait, subjonctif imparfait, subjonctif passé, subjonctif présent. We have no specific expectation regarding this variable, but we nevertheless wish to explore and control for its effect.

<sup>8</sup>Given the balance in registers, since we gathered an equal number of two registers (literature and journalism) for two constructions (COMA and SMA), this variable cannot be tested in our statistical model.



### 2.3 Data analysis

To evaluate the simultaneous effect of Event Type and Tense on the choice of verb, we used a mixed-effects logistic regression model, with Event Type as a fixed factor (4 levels: accomplishment, achievement, activity and state) and with Tense as a random intercept (19 tenses).<sup>9</sup> We examined the significance of both predictors by means of a Likelihood Ratio Test of Nested models. We interpret the effects by means of effect plots with the outcome on the probability scale. We additionally performed pairwise comparisons between the different categories of Event Type. All analyses were performed with R (R Core Team 2023) in RStudio. We used the following packages (in alphabetical order): effects (Fox 2003), ggplot2 (Wickham 2016, Wickham et al. 2021), Hmisc (Harell and Dupont 2021), lattice (Sarkar 2008), lme4 (Bates et al. 2015), lsmeans (Lenth 2016).

### 3. RESULTS

A total of  $N = 2000$  observations were annotated. Table 1 shows the bivariate frequencies and proportions of the three predictor variables under analysis in relation to the outcome variable *Verb*. The bivariate results suggest that COMA is preferably used with the Event Types of *accomplishment*, *achievement* and *state*, whereas SMA is mostly used with *activity*. For *Type of Adverb*, there were only 72 annotated observations; because of this low number, this variable was not included as a predictor variable in the multivariate model. However, based on a chi squared test of independence, strong evidence was found that the choice of verb is strongly associated with Type of Adverb ( $X^2 = 26$ ,  $df = 1$ ,  $P$ -value < 0.001, Cramér's  $V = 0.60$ ).

With regards to Tense, it is notable that both constructions occur most frequently with *indicatif imparfait*, *indicatif présent*, *passé composé*, *passé simple* and the *Infinitive mood*. Most tenses prefer COMA, but the association between SMA and the *passé simple* is a striking result to which we will return in detail in section 5. We found only a handful of observations ( $n < 5$ ) for the following tenses: *futur antérieur*, *impératif*, *infinitif passé*, *passé antérieur*, *passé récent*, *subjonctif imparfait* and *subjonctif passé*. We include these in our multivariate model, although no reliable conclusions can be drawn for these tenses, based on such a low frequency.

Next, we present the results of a mixed-effects logistic regression model with *Verb* as the outcome variable, *Event Type* as a fixed factor and *Tense* as a Random intercept. The model that we present has an average C-index of concordance of 75% (based on 200 bootstrap train and test samples), which is associated with a

---

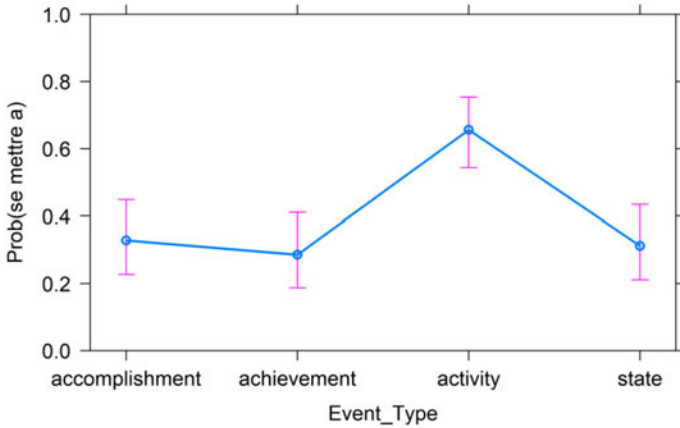
<sup>9</sup>A reviewer suggested to include lexical verb as a random factor. This is an option that we have also considered. However, upon closer inspection of the data, we observed that there were 891 different lexical verbs (on a total of 2000 observations). Of those, 786 (39%) verbs occurred only three times or less, which implies that the variability is too low to evaluate whether these lexical verbs have a specific preference for one of both constructions. Hence, we did not include lexical verb as a random factor.

	Verb	
	Commencer à(COMA)	Se mettre à(SMA)
<b>Event Type</b>		
Accomplishment	247 (68%)	116 (32%)
Achievement	159 (73%)	59 (27%)
Activity	392 (35%)	744 (65%)
State	202 (71%)	81 (29%)
<b>Adverb type</b>		
Duration and progression	19 (73%)	7 (17%)
Velocity and suddenness	5 (11%)	41 (89%)
<b>Mood and Tense</b>		
Conditionnel passé	2 (33%)	4 (67%)
Conditionnel présent	5 (31%)	11 (69%)
Futur antérieur	3 (100%)	0 (0%)
Futur proche	4 (17%)	20 (83%)
Futur simple	10 (48%)	11 (52%)
Impératif	1 (100)	0 (0%)
Indicatif imparfait	190 (68%)	89 (32%)
Indicatif présent	397 (54%)	333 (46%)
Infinitif	50 (57%)	38 (43%)
Infinitif passé	4 (100%)	0 (0%)
Participe présent	7 (58%)	5 (42%)
Passé antérieur	1 (100%)	0 (0%)
Passé composé	234 (48%)	250 (52%)
Passé récent	1 (25%)	3 (75%)
Passé simple	46 (19%)	192 (81%)
Plus-que-parfait	34 (77%)	10 (23%)
Subjonctif imparfait	0 (0%)	1 (100%)
Subjonctif passé	1 (25%)	3 (75%)
Subjonctif présent	10 (25%)	30 (75%)

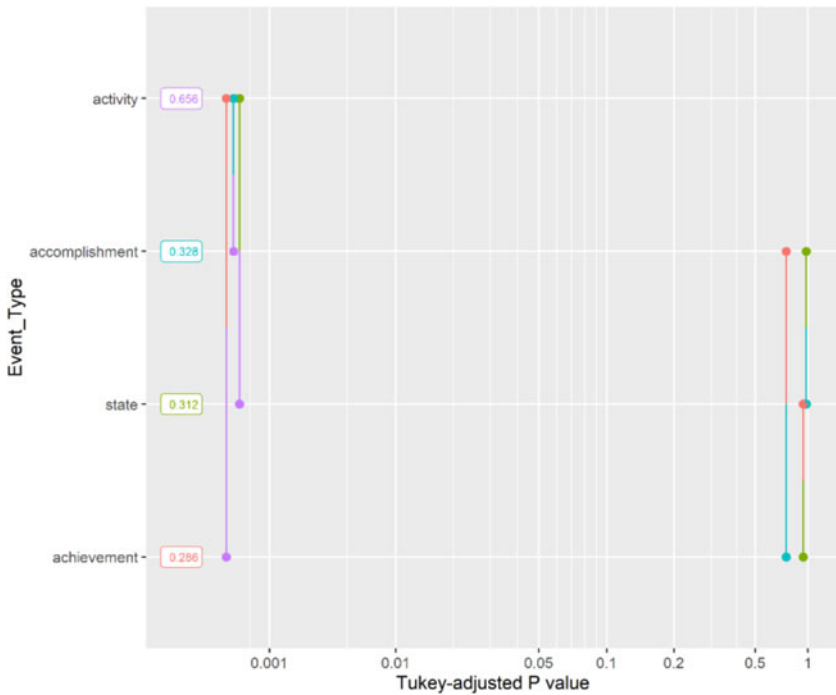
**Table 1:** Bivariate frequencies and proportions of the variables under analysis

good classification accuracy. Using the Likelihood Ratio Test of Nested models, we found evidence for both the effect of Event Type ( $G = 215$ ,  $df = 3$ ,  $P\text{-value} < 0.001$ ) and Tense ( $G = 97$ ,  $df = 1$ ,  $P\text{-Value} < 0.001$ ). The effect plot in [Figure 3](#) shows that the Event Type “activity” is most probably used with SMA, and, vice versa, that “accomplishment”, “achievement” and “state” is more likely with COMA.

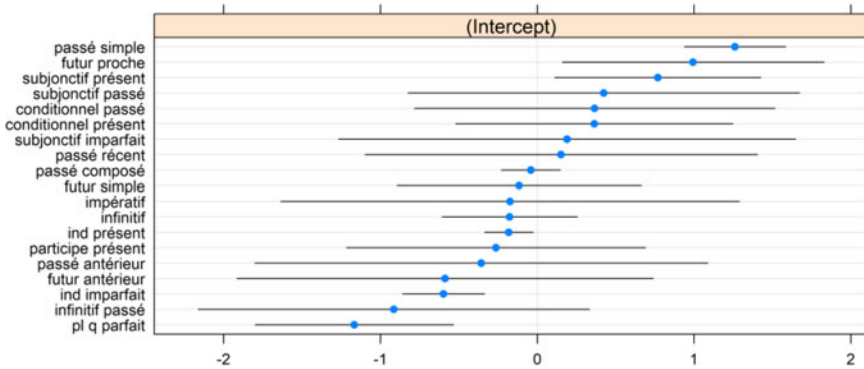
A posthoc test of the pairwise comparisons between the predicted probabilities (shown on the Y-axis in [Figure 4](#)) indicates that the difference between the activity and the three other Event Types is significant at the 5% significance level (controlling for the Family Wise Error rate with the Tukey method). No significant difference is found between accomplishment, achievement, and state. The pairwise differences can be seen in [Figure 4](#):



**Figure 3:** Effect plot showing the predicted probabilities associated with the fixed effect Event Type.



**Figure 4:** Pairwise comparisons plot. The X-axis shows the corrected P-values associated with each pairwise difference. The Y-axis shows the different Event Types together with its predicted probability for SMA.



**Figure 5:** BLUPS for the random factor (intercepts) Tense. Higher values on the X-axis (intercepts) are associated with the use of SMA, while lower values are associated with COMA.

We now turn to the result of the random factor Tense. [Figure 5](#) shows the BLUPS (Best Linear Unbiased Predictors) for the different tenses together with the 95% confidence interval:

[Figure 5](#) shows that SMA is most often used with the *passé simple*, *futur proche* and *subjonctif présent*, whereas COMA with *plus-que-parfait* and *indicatif imparfait*. Although associated with a low value on the X-axis, the *infinitif passé* only has 4 observations so the uncertainty for this BLUP is high, as indicated by the wide confidence interval (indicated by the wide line).

#### 4. QUALITATIVE ANALYSIS

This section presents a close reading of the corpus attestations. It encompasses the main results obtained in [section 3](#), elements that are systematically annotated but rejected by the multivariate model (adverb type, negation), and elements that are not systematically annotated (e.g., different types of states). The latter may offer perspectives for further research.

##### 4.1 Event type

Previous research (e.g., [Lamiroy 1987](#), [Peeters 1993](#), [Iordache and Scurtu 1994](#)) suggests a positive association between SMA and the event type activity. Our corpus findings show this association to be significant. With regard to achievements, [Iordache and Scurtu \(1994: 45\)](#) note that SMA only combines with this event type when there is a direct object that involves the idea of process (11). First, *la grandeur de l'islam* (11) does not, in our opinion, imply any idea of gradual change (process). This notion is here entailed by *Plus... plus* and the imperfect. However, a direct object expressing a process can facilitate the combination with achievements (see [Rochette 1993: 73](#)). This is illustrated in example (18). Second, we note that the direct object

does not necessarily involve an idea of process (12) and that the incompatibility is far from absolute given the frequency of achievements with the two constructions in the corpus.

- (11) Plus il taillait ses ennemis en pièce détachées (...) plus il se mettait à découvrir la grandeur de l'islam. (J.-E. Hallier cited in Iordache and Scurtu 1994: 45)  
 'The more he cut his enemies into pieces (...) the more he started to discover the greatness of Islam.'
- (12) Des compagnies pétrolières américaines qui avaient des concessions dans le sud-est du pays, où elles prétendaient n'avoir jamais trouvé de pétrole, se mettent tout à coup à en découvrir. (FT106)  
 'American oil companies that had concessions in the south-east of the country, where they claimed to have never found oil, suddenly start to discover some.'

We observe that the compatibility for both constructions can indeed depend on (i) a subject NP which is plural (13–14)<sup>10</sup> or which is a collective noun (15); (ii) an adverbial temporal determination (16); (iii) a metaphorical use of the achievement predicate (17) or (iv) the nature of the arguments of the predicate: (18) is acceptable because the direct object involves a process.

- (13) Il y a deux ans, arrêtez-moi si je me trompe, vous étiez employé comme bouc émissaire dans un grand magasin où des bombes se sont mises à exploser partout où vous passiez. (FT056)  
 'Two years ago, stop me if I'm wrong, you were employed as a scapegoat in a department store where bombs went off everywhere you went.'
- (14) Des gens commençaient à se pointer, aux alentours de la place, des lumières s'allumaient partout. (FT039)  
 'People began to show up, around the square, lights were lit everywhere.'
- (15) Il a suffi d'une pénurie alimentaire pour que la population commence à repartir vers la Tanzanie. (LM063)  
 'All it took was a food shortage for the population to begin to return to Tanzania.'
- (16) C'est sur le fond pâle de l'un de ces songes que Charlotte commença soudain à reconnaître toutes ces constellations des yeux. (FT59)  
 'It was against the pale background of one of those dreams that Charlotte suddenly began to recognize all of these constellations of the eyes.'
- (17) La lecture à chien est une méthode pédagogique qui commence à entrer dans les écoles américaines. (LM110)  
 'Reading with a dog is a teaching method that is beginning to enter American schools.'
- (18) (Elle aime tellement son amusement qu'elle commence à comprendre le jeu. (FT573)  
 'She enjoys her fun so much that she begins to understand the game.'

<sup>10</sup>According to Iordache and Scurtu (1994: 45), the interpretation of iterative action is only possible with COMA. Example (13) contradicts this assertion.

However, these conditions do not cover all of the contexts where the combination with achievements is possible:

- (19) Poussé par le public, il s'est mis à shooter. Et à marquer. (LM077)  
'Pushed by the spectators, he started shooting. And scoring.'
- (20) Et d'autres assassins potentiels seront de fait encouragés si, pour limiter la liberté d'expression, on se met à invoquer l'éthique de responsabilité. (LM079)  
'And other potential assassins will in fact be encouraged if, to limit the freedom of expression, one starts to invoke the ethics of responsibility.'

Our corpus data show, first, for the combination with states, that the constraint for SMA is less strong than previous research suggests. In addition, the stative constraint for both constructions should be relativized, as the number of states is found to be more frequent than that of achievements.

- (21) Les Français se sont mis à préférer boire chez eux. (LM065)  
'The French have started to prefer drinking at home.'
- (22) Des bougies commencent à briller sur les tables... (FT048)  
'Candles begin to shine on the tables ...'
- (23) La France en est là : obsédée de crimes sexuels. La machine judiciaire en est gavée. (...) Mais le sexe commence, lui aussi, à lasser. On s'intéresse désormais davantage à l'école. (LM075)  
'France is there: obsessed with sex crimes. The judicial machine is force-fed. (...) But people are beginning to tire of it. They are now more interested in school.'
- (24) On commence à posséder un passé, comme on possède une maison. (FT055)  
'One begins to own a past, as one owns a house.'
- (25) Le même phénomène commence à s'observer avec Diego, le cousin de Dora, lui aussi en voie de fonder un empire avec son émission, lancée récemment. (LM076)  
'The same phenomenon begins to be observed with Diego, Dora's cousin, He also is in the process of building an empire with his recently launched show.'
- (26) Son nom se mit à être connu au-delà du cercle de ses familiers. (FT051)  
'His name came to be known beyond the circle of his close circle.'

Peeters (1993) claims that the stative constraint for SMA disappears in certain contexts, more precisely with an iterative reading, or with another verb intervening, or if there is a clear action element in glosses with a state predicate. However, these contexts do not convince us. As for the contexts proposed by Lamiroy (1987), we actually see the use of a subject NP in the plural (21)–(22), the use of generic NP (23)–(24), of the *se-moyen* (25) and adverbial complements (26). However, we find that these syntactic situations, though they contribute to a sequential interpretation, are not sufficient to explain the following utterances.

- (27) L'avance des chars allemands trace son chemin de Damas dans le plat pays flamand : il se met à haïr ce qu'il adorait et à aimer ce qu'il détestait. (FT042)  
'The advance of the German tanks traces his road to Damascus through the flat Flemish country: he starts to hate what he adored and to love what he hated.'
- (28) Ce n'est pas parce que tu te mets à avoir un cheveu blanc... (FT052)  
'It's not because you start to have white hair...'

- (29) Et puis, Julien commençait à me manquer sérieusement. (FT054)  
 ‘Besides, I was beginning to seriously miss Julien.’

We therefore observe that the stative constraint is present in some contexts, but not in others (27)–(29). Moreover, the constraint is different depending on the ingressive construction. Examples (30)–(34), possible with COMA, show that SMA is more sensitive to the stative constraint.

- (30) \*Claude se met à posséder une belle collection de boîtes de camembert. (Saunier 1999: 276)  
 ‘Claude starts owning a fine collection of Camembert boxes.’
- (31) \*Tout le monde se met à le savoir, que tu as réussi tes examens. (Saunier 1999: 276)  
 ‘Everyone starts to know that you passed your exams.’
- (32) \*Jean se met à être capitaine. (Lamiroy 1987: 293)  
 ‘Jean starts being a captain.’
- (33) \*Josyane s’est mise à sentir des douleurs hier. (Schmid 1984: 56)  
 ‘Josyane started to feel pain yesterday.’
- (34) \*Les fleurs se mirent à être agréables à voir. (Peeters 1993: 40)  
 ‘The flowers started to be nice to see.’

The observations of Peeters (1993) and Saunier (1999) already indicate that SMA selects state predicates in a more restricted way. Our corpus confirms that SMA prefers some states to others. As a result, several explanations can be formulated as to why state compatibility presents itself differently for SMA and for COMA. Saunier (1999: 276) remarks that one finds *aimer* ‘to love’ more easily with SMA than with COMA and that it is therefore not a question of constraints linked to the aspectual properties of the event, but to the way in which the subject’s relation to the event can be constructed. The author emphasizes that the potential exteriority of loving and knowing in relation to a subject is not identical and that it is precisely factors of incongruity and suddenness that take precedence over the purely aspectual question. These factors do not really stand out from our examples. Another hypothesis to be taken into account could be the different types of states, namely permanent states vs. stage-level/transitory states. At first glance, SMA combines more with transitory states (27) than with permanent states (30)–(31). However, we observe that permanent states cannot be excluded (28), while certain transitory states are not allowed (33)–(34). We find neither of these hypotheses to be convincing. The only thing that is certain, however, is that state compatibility looks different for SMA than for COMA. Marque-Pucheu (1999: 240) notes that SMA + state predicate corresponds to a property difference, namely the possibility of adding the adverb of degree *encore plus* ‘even more’. Indeed, we observe that this adverb can be added to states (32)–(34), which makes (35)–(37) more acceptable. However, example (31), repeated here as (38), remains problematic.

- (35) Jean se met à être encore plus capitaine qu’avant.  
 ‘Jean starts to be even more of a captain than before.’
- (36) Josyane s’est mise à sentir encore plus de douleurs hier.  
 ‘Josyane started to feel even more pain yesterday.’

- (37) Les fleurs se mirent à être encore plus agréables à voir.  
‘The flowers started to look even nicer.’
- (38) \*Tout le monde se met à le savoir encore plus, que tu as réussi tes examens.  
‘Everyone starts to know, even more, that you have passed your exams.’

If the two ingressive constructions can be made compatible with permanent or transitory states thanks to certain syntactic conditions, it is appropriate to add a further syntactic condition for SMA, namely that of the progression (35)–(37) which COMA already includes in its semantics. This progression reminds us of the example of Franckel (1989: 145) *Je commence à en avoir marre* ‘I’m starting to get fed up’ that he interprets as *Je commence à en avoir vraiment marre* ‘I’m starting to get really fed up’, which cannot be replaced by ??*Je me mets à en avoir marre*. It can indeed be said that, for states, *commencer à P* denotes the transition from *not really P* to *really P*, while *se mettre à P* only indicates the transition from *not P* to *P*, hence the suddenness effect associated with SMA, while COMA rather profiles a gradual inception.

## 4.2 Mood and Tense

We find that the Simple Future tense is observed both with SMA and with COMA. SMA is likely to appear with all the tenses and moods, but the realizations in the Imperative, the Future Perfect and the Past Anterior are lacking in our corpus. The results show clear differences between the two ingressive constructions in relation to the selected tenses and moods. It is clear that the SMA construction has a strong preference for the Simple Past in both corpora. The Simple Past translates the perfective aspect which is bounded and internally heterogeneous, in contrast to the imperfective which is unbounded and internally homogeneous. We believe that COMA behaves in a neutral way with respect to different tenses and moods. On the other hand, the remarkable frequency of the Simple Past with SMA does not seem accidental to us. What the semantics of SMA shares with the perfective aspect conveyed by the Simple Past is its punctual character, in other words, the beginning of an event is not perceived in an incremental way.

## 4.3 Adverb type and negation

Several examples of SMA with adverbs of velocity and suddenness can be found in our corpus.<sup>11</sup>

- (39) Aussitôt, le chevalier de Beltram se mit à rougir comme un enfant. (FT043)  
‘Immediately the knight of Beltram started to blush like a child.’
- (40) L’homme qui m’interroge avec des yeux cruels se met brusquement à tousser. (FT044)  
‘The man who questions me with cruel eyes suddenly starts to cough.’

One might wonder whether it is really the scope of the adverbs of suddenness that is in question here, or whether it is the perfective value generated by the Simple

<sup>11</sup>See also the use of *tout à coup* in example (12).



Past (39). Example (40) shows that this combination is also possible with the Simple Present. Sato (1994) goes so far as to suggest that the distribution of this type of adverbs is exclusively reserved for SMA, in other words that they cannot appear with COMA. However, example (16), repeated here as (41), contradicts this assertion.

- (41) C'est sur le fond pâle de l'un de ces songes que Charlotte commença soudain à reconnaître toutes ces constellations des yeux. (FT059)  
 'It was against the pale background of one of those dreams that Charlotte suddenly began to recognize all of these constellations of the eyes.'

Finally, we note with Peeters (1993) and Saunier (1999) that the use of adverbs of suddenness is not systematic and that opposite evidence exists, as in (42).

- (42) Il se mettait lentement à comprendre qu'à un certain niveau de la finance et de la politique américaine les juifs, si extraordinairement commodes par leur agilité intellectuelle dans les tâches subalternes, devenaient très vite un peu gênants. (FT045)  
 'He was slowly coming to realize that at a certain level of American finance and politics the Jews, so extraordinarily convenient for their intellectual agility in menial tasks, very quickly became a bit of a nuisance.'

Table 2 shows that the individual numbers of adverbs are generally too small to draw serious conclusions. Nevertheless, we see that the frequency of this kind of adverb is still higher for SMA. Instead of saying that SMA marks a more "brutal" inception than COMA (Franckel 1989: 147), we prefer to say that SMA marks the initial phase in a punctual way.

As for the adverbs of duration and progression, we mainly consider the phasal use<sup>12</sup> of the adverb *déjà* 'already' which indicates that the event is started earlier than expected. Mosegaard Hansen and Strudsholm (2008: 477) note that, in the phasal use, *déjà* is compatible primarily with states and activities and will not normally occur with the perfective past. Since SMA and COMA mark durative events, they occur with this adverb (43)–(44).

- (43) Au moins par l'assurance, l'autorité, le sans-gêne, il se mettait déjà à ressembler à ces hommes dont, quelques heures plus tôt, il ne savait encore rien. (FT046)  
 'At least by the assurance, the authority, the shamelessness, he was already starting to look like those men of whom, a few hours earlier, he still knew nothing.'
- (44) Les hommes de main, serviles, commencent déjà à rire tandis que le potentat cherche nonchalamment dans son imagination le plus raffiné des supplices. (FT047)  
 'The henchmen, servile, are already beginning to laugh while the potentate casually searches for the most refined of tortures in his imagination.'

<sup>12</sup>One of the numerous uses of *déjà* described by Mosegaard Hansen and Strudsholm (2008) and which they define (p. 477) as: "phasal *déjà/già* assert that the state-of-affairs (SoA) *e* expressed in their host sentence has begun prior to a topic time TT, understood as the time with respect to which the main claim of the utterance is made [...], that is, the initial phase of *e* has occurred at TT."

	SMA	COMA
<i>aussitôt</i>	6	0
<i>brusquement</i>	2	0
<i>du coup</i>	1	1
<i>d'un coup</i>	2	0
<i>immédiatement</i>	1	0
<i>sitôt</i>	0	1
<i>soudain</i>	11	1
<i>subitement</i>	1	0
<i>tout à coup</i>	8	1
<i>tout de suite</i>	2	0
<i>tout d'un coup</i>	3	0
<i>vite</i>	4	1
<b>Total</b>	<b>41</b>	<b>5</b>

**Table 2:** Distribution of adverbs of velocity and suddenness

	SMA	COMA
<i>de plus en plus</i>	1	1
<i>déjà</i>	2	16
<i>petit à petit</i>	1	0
<i>peu à peu</i>	3	1
<i>progressivement</i>	0	1
<b>Total</b>	<b>7</b>	<b>19</b>

**Table 3:** Distribution of adverbs of duration and progression

However, we note (Table 3) that COMA clearly appears more frequently with *déjà* than with SMA. Franckel (1989: 144) notes that one of COMA's essential characteristics is to involve a first construction of P (process), through anticipation. Instead of "anticipation to this first construction of P", we prefer saying that COMA implies a first phase of P. The adverbs *de plus en plus* 'more and more' and *progressivement* 'gradually' converge with the idea of a first construction of P through a progression. While SMA prefers selecting the sub-event punctually, COMA prefers marking this first construction of an event gradually.

With regard to negation, Franckel (1989: 144) states that the negative form is very strongly constrained, if not impossible, with SMA. However, several examples in the data do not support this hypothesis.

- (45) Jamais cette masse humaine ne se met à gronder : « Pourquoi eux et pas nous? » (LM072)  
 'This human mass never starts to scold: "Why them and not us?"
- (46) Je ne vais quand même pas me mettre à pleurer! (LM074)  
 'I'm certainly not going to start crying!'

Franckel (1989: 144) adds that the negation of SMA has no status, due to the aoristic functioning of this construction. Although the difference in negative modality does not translate quantitatively in our corpus, we still observe that the initial phase for SMA is taken as a whole and that the negation consequently eliminates the existence of the entire phase. On the other hand, the negation relating to COMA is often only a relative negation (47) *pas encore* ‘not yet’, (48) *jamais vraiment* ‘never really’ and (49) where one can easily add *encore* ‘again’.

- (47) Ces dernières n’ont pourtant pas encore commencé à les publier (LM069)  
 ‘The latter have however not yet begun to publish them.’
- (48) D’une certaine façon, la nouvelle présidente chilienne incarne le modernisme réformateur que le PS paraît appeler de ses vœux mais dont il n’a jamais vraiment commencé à décrire les contours. (LM070)  
 ‘In a way, the new Chilean president embodies reformist modernism that the PS seems to be calling for but whose outlines it has never really begun to describe.’
- (49) Le fait que cette manipulation concerne des centrifugeuses isolées, et non reliées en cascade, signifie que l’Iran n’a pas commencé à procéder à un enrichissement d’uranium à grande échelle. (LM071)  
 ‘The fact that this manipulation concerns isolated centrifuges, rather than a cascade of centrifuges, means Iran has not begun enriching uranium on a large scale.’

This difference is also evident with the adverb *à peine* ‘barely’ (SMA N = 0; COMA N = 6). The use of *à peine* with the SMA construction indicates that the event has not (yet) taken place. On the other hand, with COMA, the event has been started, and its first construction is highlighted. This aspectual interpretation of COMA + *à peine* is not really more appealing than that of Vallet,<sup>13</sup> who pronounces on the qualitative value of the combination. The alleged imperfection noted in her example *Pierre commence à peine à lire* ‘Pierre is just beginning to read’ cannot however be generalized, as shown in (50)–(51).

- (50) Les arbres des trois îlots situés un peu plus en aval commençaient à peine à bourgeonner. (FT062)  
 ‘The trees of the three islets located a little further downstream were just beginning to bud.’
- (51) La demeure a été transformée en foyer d’accueil pour tout jeunes rescapés. Ces enfants juifs ont survécu de justesse. Ils commencent à peine à revivre. (LM082)  
 ‘The house has been transformed into a foster home for very young survivors. These Jewish children barely survived. They are just beginning to come back to life.’

<sup>13</sup>Vallet (2006); « Les deux constructions [*se mettre à* et *commencer à*] s’opposent aussi sous le rapport des adverbes semi négatifs qu’elles acceptent ou refusent : dans *Vous ne me dérangez pas, je m’étais à peine mis à lire*, la locution *à peine* nie l’incohation (*je n’avais pas vraiment commencé à lire*) et non la manière de lire, en revanche dans *Pierre (commence + \*se met) à peine à lire*, l’expression *commencer à* n’est pas un argument pour accomplir parfaitement et *à peine* confirme cette imperfection (*Pierre peine à lire, Pierre ne lit pas vraiment bien*). »

In sum, the constraint of adverbial complementation is not absolute. Although the data is limited, we can still distinguish some trends: COMA prefers to approach the sub-event more gradually and less overall, while SMA prefers to approach the sub-event immediately and globally.

## 5. FRAME-SEMANTIC REPRESENTATION

We now proceed to interpret the corpus-based results from a frame-semantic perspective (Croft 2012). Our observations in relation to the event types, the adverbial complementation and the verbal tenses suggest that the two constructions yield a different construal of the initial phase of the event expressed. With SMA, the beginning of the event is expressed as something punctual, while with COMA, the first phase of the event is presented more gradually. This last point is present in statements where the event is interrupted shortly after having started (52).

- (52) a. J'avais commencé à dessiner mes personnages, et il a fallu tout modifier.  
 b. ??Je m'étais mise à dessiner mes personnages, et il a fallu tout modifier.  
 (Saunier 1999: 285)  
 'I had started to draw my characters, and everything had to be changed.'

When the event is in this first phase, it is often necessary, with COMA, to mark the true trigger with an adverb like *vraiment*, *sérieusement*, etc. 'really, seriously' (53)–(54). However, there is no evidence of this phenomenon for SMA in our corpus. This absence is explained by the fact that SMA captures the beginning of the event in a global way, regardless of its internal structure.

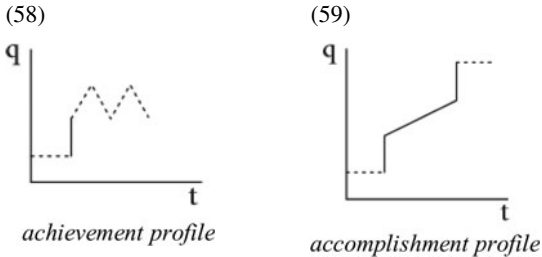
- (53) La nuit commençait à vraiment tomber. (FT067)  
 'Night was really beginning to fall.'  
 (54) En fait, c'est quand tes oreilles s'y mettent que ça commence à sérieusement commencer. (FT069)  
 'In fact, it's when your ears get started that it definitely begins.'

The combination with the Progressive *être en train* seems difficult for SMA. However, unlike what Laca (2002) claims (55), it is possible for COMA (56)–(57). This suggests that COMA allows a durative reading of the initial phase of the event.

- (55) ??Il était en train de commencer à avoir des doutes. (Laca 2002: 89)  
 'He was beginning to have doubts.'  
 (56) Ma vraie vie de chatte, c'était de me prélasser sans fin dans ce lit, ma vraie vie de chatte, j'étais en train de tout juste commencer à la vivre. (FT070)  
 'My real cat life was endless lounging in this bed, my real cat life, I was just beginning to live it.'  
 (57) C'est comme si on faisait une séance d'exorcisme, là, Liv et moi, Liv montée sur moi, tendue en arrière, avec ses seins dans ses mains, en train de commencer à jouir, mais non, pas tout de suite, patience. (FT071)  
 'It's like having an exorcism session there, Liv and I, Liv riding

on top of me, stretched backwards, with her breasts in her hands, beginning to come, but no, not right away, patience.’

Both ingressive constructions mark the initial transitional phase, but they modify the aspect contour of a base event in different ways. COMA is able to exhibit an achievement profile (58)<sup>14</sup> or accomplishment profile (59)<sup>15</sup> while SMA tends to mark a punctual transition (58). In other words, COMA, unlike SMA, has the potential to profile the initial phase in two distinct ways.<sup>16</sup>



Aspectual adverbs and verbal tenses join the semantics of each construction. As for the basic events on which our constructions operate, we have noticed that SMA favors activities. A distinctive collexemic analysis (Verroens 2011) has shown that the significant collexemes that come into play for SMA are essentially part of the semantic classes of non-verbal (*crying, laughing, whining*) or verbal (*shouting, yelling, speaking*) expression, acts of performance (*dancing, playing*) and movement (*running, moving, pushing, turning*). An examination of collexemes statistically rejected for SMA, but significant for COMA, reveals the overall meaning of the SMA construction, namely the punctual sense. This meaning is intrinsic, or more precisely, it is not inferred by collexemes. The punctual value of SMA is not only compatible with expressive collexemes, but, moreover, emphasizes their unforeseen and unintentional meaning. In contrast, the incremental value of COMA is consistent with

<sup>14</sup>More precisely, Croft (2012) speaks of a directed achievement profile in (58), and an incremental accomplishment in (59). This elaborated terminology (see section 1), corresponding to the Vendlerian categories of achievement and accomplishment respectively, is less important here.

<sup>15</sup>In (59), COMA presents the profile of an accomplishment: an inceptive transition phase, a directed change phase and a completive transition phase, just like an accomplishment. However, the unprofiled part after the completive transition phase would have the aspectual contour appropriate to that event, which could be something other than the stative phase shown in (59).

<sup>16</sup>See Laca (2004: 86) who notes that the ingressive construction alters the basic event. For instance, in *Le malade commençait à s'affaiblir* 'The sick person began to weaken' COMA modifies the accomplishment of the predicate *s'affaiblir* to render an achievement that corresponds to the initial transition of that accomplishment. In our opinion, COMA presents in this example rather a gradual initial transition and thus aligns with the event type of the basic event (weakening).

the gradual onset of cognitive and sensory collexemes. We hypothesize that the semantic distinction between the two constructions is the result of a linguistic change. An in-depth diachronic study, outside the scope of this article, would need to show exactly what kind of change it is, but, for the time being, there are already indications that COMA + Vinf. became more durative in the 13th century (Beer 1974), shortly after the first attestation of SMA + Vinf. at the end of the 12<sup>th</sup> century.

What are, then, the consequences of our results for phasal aspect and the overall concept of aspect? With respect to phasal aspect, we were able to observe that each construction marks the sub-event (i.e., the initial phase of the event) in a different way: COMA builds a sub-event corresponding to the initial phase of the event (designated by the infinitive), and this non-punctual initial phase can just as easily be grasped under a perfective as an imperfective construal (*commença à/ était en train de commencer à* ‘began to / was beginning to), whereas SMA rather designates the initial boundary of the event. SMA thus constructs a punctual phase, which, because of its punctuality, is very difficult to reconcile with the imperfective, but is perfectly compatible with the perfective and with punctual adverbs. This result begs the question of how lexical and grammatical aspect interact. The data clearly show that there is a strong preferential (prototypical) association between congruent semantic values. On the one hand, we have telicity and perfectivity expressed by SMA (the adverbs of suddenness and the Simple Past, respectively) and, on the other hand, the unbounded character of COMA (the adverbs of duration and the imperfect).<sup>17</sup> In the structure [<sub>TP</sub> Tense [<sub>TP</sub> Aspect [<sub>VP</sub> Aspectual class]]], the ingressesives are perceived as lexical items (COV-MAP, see section 1) which modify the aspectual class by adding the features [INGR, + TEL ] to the VP. The Simple Past, on the other hand, is considered as a grammatical marker which not only brings in TP<sub>tense</sub> the feature [PAST] and in TP<sub>aspect</sub> the feature [+ PERF], but which is also able to convey the property [INGR] when it is combined with non-punctual predicates, in particular states (60), but also activities (61).

(60) Elle aime les romans de Walter Scott. (Riegel et al. 1994: 304)  
‘She began to like the novels of Walter Scott.’

(61) Elle entra dans la chambre funéraire et à la vue de sa mère défunte, elle pleura. (Ciszewska 2004: 9).  
‘She entered the burial chamber and at the sight of her deceased mother, she started to cry.’

The ingressive meaning is not an intrinsic property of the Simple Past, but rather a pragmatic-discursive characteristic depending on the context. The inceptive interpretation appears when the basic reading is implausible. In (60), the unbounded

<sup>17</sup>In this sense, our results confirm the general hypothesis that there is an intimate relationship between perfectivity and telicity or punctuality on the one hand and imperfectivity and stativity on the other. The first affinity is often interpreted in terms of “boundedness”, while the second is interpreted in terms of “unboundedness” (Sasse 2002: 205). Unlike telicity, perfectivity (the boundedness imposed by the Simple Past) is not intrinsic but grammatically established.

event (liking Walter Scott's novels) is grounded in favor of a profiled event which is bounded, ingressive and made possible by the Simple Past. In (61), the punctual modifier *à la vue de sa mère défunte* leads to an inceptive reading of the main clause.<sup>18</sup> So there seems to be a strong convergence between the lexical and grammatical aspect, in particular between SMA and the Simple Past. This suggests that the interpretation of the Simple Past is closely linked to the semantics of the verb to which it is applied, instead of being completely independent of it, as is the case in the two-component approach (section 1). Indeed, the adherents of the one-component approach claim that it is precisely the affinity between (non-) telicity and (non-) perfectivity that their theory captures in terms of a single set of primitives. According to Croft (2012: 127–172), the perfective/imperfective distinction in the grammatical aspect corresponds to an opposition of aspectual interpretations within the lexical aspect.<sup>19</sup> His conclusion is based on the typological study by Bybee et al. (1994) carried out on 90 languages and on a multidimensional scaling (MDS) carried out by Croft and Poole (2008) on data from Dahl (1985). The distribution of the lexical aspect resulting from the MDS analysis suggests that the punctual events (achievements) are closest to the perfective ending in the aspectual dimension. With regard to our data, we come to the same conclusion although we have used other multivariate techniques (section 3): the event type that one obtains when SMA alters the main verb, namely an achievement, is linked with the perfective while COMA, in case of an accomplishment, is connected to the imperfective.

## 6. CONCLUSIONS

On a descriptive level, we can conclude that most of the intuitive observations found in the literature are justified. Our corpus findings suggest that SMA is significantly associated with activities. SMA appears to be associated with *passé simple*, *futur proche* and *subjonctif présent*, whereas COMA is associated with *plus-que-parfait* and *indicatif imparfait*. As for the adverbial complementation and the somewhat impressionistic assumption of some linguists about the “brutality” of the inception, we have no evidence that this is indeed the case, due to the limited number of instances found with an adverbial complement. Nevertheless, the results indicate that SMA builds a punctual sub-event that aligns with the perfective and with punctual adverbs. Lastly, the qualitative analysis has shown that the combination of ingressive constructions with inherent permanent states is not totally impossible, contrary to what is often claimed in the literature.

<sup>18</sup>For a detailed analysis of aspectual coercion, see in particular de Swart (1998), Michaelis (1998, 2004) and Croft (2012).

<sup>19</sup>Croft (2012: 171–172): “Our conclusion is that the perfective/imperfective grammatical opposition represents two families of closely related aspectual construals that are characteristic of opposite sides of the lexical aspect circle [...]. The opposing construals are grounded not in a single binary feature opposition, but on opposing aspectual contours, the directed contour for perfective and the undirected contour for imperfective.” Croft argues that in this way the spatial model of lexical aspect provides a new basis for a semantic definition of grammatical aspect.

On a theoretical level, we have used Croft's (2012) spatial model of lexical aspect, which argues for a usage-based, encyclopedic semantic model, whereby frequency of use determines the strength of the alternative construals and whether or not there is a default. Croft's model has clear potential for the description of aspectual constructions. We have argued that, as for predicates (see the example *to see* illustrated in section 1), aspectual constructions can also have more than one construal. The two ingressive constructions can be distinguished on the basis of clear aspectual differences in terms of a punctual/durative analysis. COMA can have an achievement profile or that of accomplishment, while SMA manifests only one profile, more precisely the achievement profile. Finally, our results support a one-component approach to aspect. Although we have used other statistics, and although the scope of our study is limited to French ingressives, we come to the same conclusion: grammatical aspect interacts with lexical aspect, but the result is of the same semantic type (aspectual contour) as lexical aspect.

#### REFERENCES

- Barceló, Gérard Joan, and Jacques Bres. 2006. *Les temps de l'indicatif en français*. Paris: Ophrys.
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Beer, Jenette. 1974. Traces of a durative use of *comencier* + infinitive in 13th-century French. *Romance Philology* 28(1): 43–48.
- Bertinetto, Pier Marco, and Denis Delfitto. 2000. Aspect vs. actionality: Why they should be kept apart. In *Tense and aspect in the languages of Europe*, ed. Ö. Dahl, 189–226. Berlin: de Gruyter.
- Bybee, Joan, Revere Perkins, and William Pagliuca, eds. 1994. *The evolution of grammar: Tense, Aspect and Modality in the languages of the world*. Chicago: University of Chicago Press.
- Cipria, Alicia, and Craige Roberts. 2000. Spanish *imperfecto* and *pretérito*: Truth conditions and Aktionsart effects in a situation semantics. *Natural Language Semantics* 8(4): 297–347.
- Ciszewska, Ewa. 2004. L'inchoatif et les moyens de son expression en français. *Neophilologica* 16: 7–19.
- Coseriu, Eugenio. 1976. *Das romanische Verbalsystem* [The Romance verbal system]. Tübingen: TBL Verlag Gunter Narr.
- Croft, William. 2009. Aspectual and causal structure in event representations. In *Routes to language development. Studies in honor of Melissa Bowerman*, ed. V. C. Mueller Gathercole, 139–166. Mahwah, NJ: Erlbaum.
- Croft, William. 2012. *Verbs: Aspect and causal structure*. Oxford: Oxford University Press.
- Croft, William, and Keith T. Poole. 2008. Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics* 34(1): 1–37. doi:10.1515/thli.2008.001
- Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.
- Depraetere, Ilse. 1995. On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy* 18(1): 1–19.



- Franckel, Jean-Jacques. 1989. *Étude de quelques marqueurs aspectuels du français*, Langue et cultures 21. Genève: Droz.
- Filip, Hana. 1999. *Aspect, situation types and noun phrase semantics*. New York: Garland.
- François, Jacques. 2003. *La prédication verbale et les cadres prédictifs*. Louvain: Peeters.
- Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(1): 1–27. <https://doi.org/10.18637/jss.v008.i15>.
- Gaotone, David. 1995. Syntaxe et sémantique: Le cas des verbes ‘transparentes’. *Perspectives. Revue de l’université hébraïque de Jérusalem* 2: 55–71.
- Gardiès, Jean-Louis. 1981. Éléments pour une grammaire de l’aspect. *Modèles linguistique* 3: 112–134.
- Gosselin, Laurent. 2010a. Les relations entre périphrases aspectuelles et conjugaisons en français. In *Interpréter les temps verbaux*, ed. Nelly Flaux, Dejan Stosic, and Co Vet. Bern: Peter Lang Verlag.
- Gosselin, Laurent. 2010b. Entre lexicque et grammaire : Les périphrases aspectuelles du français. *Cahiers de lexicologie* 96: 67–95.
- Gosselin, Laurent. 2011. L’aspect de phase en français: Le rôle des périphrases verbales. *Journal of French Language Studies* 21(2): 149–171.
- Gosselin, Laurent. 2020. *Les périphrases aspectuelles du français*, in Encyclopédie grammaticale du Français. <http://encyclogram.fr>.
- Gross, Maurice. 1975. *Méthodes en syntaxe : Régime des constructions complétives*. Paris: Hermann.
- Harell, Frank E., and Charles Dupont. 2021. *Hmisc: Harrell Miscellaneous* (4.5-0). <https://CRAN.R-project.org/package=Hmisc>.
- Imbs, Paul. 1960. *L’emploi des temps verbaux en français moderne : Essai de grammaire descriptive*. Paris: Klincksieck.
- Iordache, Emilia-Rodica, and Gabriela Scurtu. 1994. Étude sémantique et syntaxique des périphrases verbales marquant le début d’accomplissement du procès. *Cahiers de Linguistique Théorique et Appliquée* XXXI: 41–48.
- Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic: Introduction to modeltheoretic semantics, formal logic and discourse representation theory*. Dordrecht: Kluwer.
- Kronning, Hans. 2003. Auxiliarité, énonciation et rhématicité. Dans *Modes de repérages temporels*, ed. S. Mellet et V. Marcel : 231–249. Amsterdam: Rodopi.
- Laca, Brenda. 2002. Spanish “aspectual” periphrases: Ordering constraints and the distinction between situation and viewpoint aspect. In *From words to discourse. Trends in Spanish Semantics and Pragmatics*, ed. Javier Gutiérrez-Rexach. Oxford: Elsevier.
- Laca, Brenda. 2004. Les catégories aspectuelles à expression périphrastique : Une interprétation des apparentes lacunes du français. *Langue française* 141: 85–98.
- Laca, Brenda. 2005. Périphrases aspectuelles et temps grammatical dans les langues romanes. In *Les périphrases verbales*, ed. Hava Bat-Zeev Shyldkrot and Nicole Le Querler. *Lingvisticae Investigationes. Supplementa* 25. Amsterdam: John Benjamins.
- Lamiroy, Béatrice. 1987. The complementation of aspectual verbs in French. *Language* 63(2): 278–298.
- Lamiroy, Béatrice. 1999. Auxiliaires, langues romanes et grammaticalisation. *Langages* 135: 33–45.
- Lenth, Russell V. 2016. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software* 69(1): 1–33. <https://doi.org/10.18637/jss.v069.i01>.
- Marque-Pucheu, Christiane. 1999. L’inchoatif : Marques formelles et lexicales et interprétation logique. In *La modalité sous tous ses aspects*, ed. Svetlana Vogeleeer, Andrée Borillo, Marcel Vuillaume, and Carl Veters, 233–257. Amsterdam: Rodopi.

- Martin, Robert. 1971. *Temps et aspect : Essai sur l'emploi des temps narratifs en moyen français*. Paris: Klincksieck.
- Michaelis, Laura A. 1998. *Aspectual grammar and past time reference*. London: Routledge.
- Michaelis, Laura A. 2004. Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics* 15(1): 1–67
- Moens, Marc, and Mark Steedman. 1988. Temporal ontology and temporal reference. In *Computational Linguistics* 14(2): 15–28.
- Mosegaard Hansen, Maj-Britt, and Erling Strudsholm. 2008. The semantics of particles: Advantages of a contrastive and panchronic approach: A study of the polysemy of French *déjà* and Italian *già*. *Linguistics* 46(3): 471–505. doi:10.1515/ling.2008.016
- Nef, Frédéric. 1980. Les verbes aspectuels du français : Remarques sémantiques et esquisse d'un traitement formel. *Semantikos* 4(1): 11–46.
- Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge MA: MIT Press.
- Peeters, Bert. 1993. *Commencer et se mettre à* : Une description axiologique-conceptuelle. *Langue française* 98: 24–47.
- R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Riegel, Martin, Jean-Christophe Pellat, and René Rioul. 1994. *Grammaire méthodique du français*. Paris: Presses Universitaires de France.
- Rochette, Anne. 1993. À propos des restrictions de sélection de type aspectuel dans les complétives infinitives du français. *Langue française* 100: 67–82.
- Roy, Gérard-Raymond. 1976. *Contribution à l'analyse du syntagme verbal : étude morpho-syntaxique et statistique des coverbes*. Québec: Presses de l'Université de Laval.
- Sarkar, Deepayan. 2008. *Lattice: Multivariate data visualization with R*. Berlin: Springer-Verlag. <https://doi.org/10.1007/978-0-387-75969-2>.
- Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6(2): 199–271.
- Sato, J. 1994. Valeurs sémantiques de *se mettre à* et *commencer à*. *Bulletin d'études de linguistique française* 28: 30–35.
- Saunier, Évelyne. 1999. Contribution à une étude de l'inchoation : *se mettre à* + inf. : Contraintes d'emploi, effets de sens et propriétés du verbe *mettre*. In *La modalité sous tous ses aspects*, ed. Svetlana. Vogeeler, Andrée Borillo, Marcel Vuillaume, and Carl Veters, 259–288. Amsterdam: Rodopi.
- Schmid, Annemarie. 1984. *Ein Beitrag zur Phraseologie des Französischen; Syntaktisch-semantiche Untersuchung zu mettre in festen Verbindungen* [A contribution to the phraseology of French: Syntactico-semantic investigation of *mettre* in solid compounds]. Innsbruck: Amoe.
- Smith, Carlota. 1997 [1991]. *The parameter of aspect* (2nd edition). Dordrecht: Kluwer academic.
- Squartini, Mario. 1998. *Verbal periphrases in Romance: Aspect, actionality, and grammaticalization*. Berlin: Mouton de Gruyter.
- de Swart, Henriëtte. 1998. Aspect shift and coercion. *Natural language and linguistic theory* 16(2): 347–385.
- de Swart, Henriëtte Elizabeth. 2012. Verbal aspect across languages. In *The Oxford handbook of Tense and Aspect*, ed. Robert I. Binninck, 752–780. Oxford: Oxford University Press.
- Vallet, Catherine. 2006. Agentivité et homogénéité dans les constructions verbales. In *Constructions verbales et production de sens: Actes du colloque organisé à Besançon, les 26, 27 et 28 janvier 2006*, ed. Daniel Lebaud, Catherine Paulin, and Katia Ploog, 253–266. Besançon: Presses Universitaires de Franche-Comté.

- Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, New York: Cornell University Press.
- Verbert, Christ'l. 1979. Description sémantico-syntaxique de *commencer*. *Travaux de linguistique* 6: 57–81.
- Verkuyl, Henk. 1999. *Aspectual issues*. Stanford: CSLI.
- Verroens, Filip. 2011. La construction inchoative *se mettre* à: Syntaxe, sémantique et grammaticalisation. Doctoral dissertation. Ghent University.
- Verroens, Filip. 2018. Sur la notion d'inchoatif en linguistique française. *Travaux de linguistique* 76(1): 91–111.
- Vet, Co. 2008. Six traits sémantiques suffisent à décrire tous les temps du français. In *L'énonciation dans tous ses états. Mélanges offerts à H. Nølk à l'occasion de ses soixante ans*, ed. Merete Birkelund, Maj-Brit Mosegaard-Hansen, and Coco Norén, 451–471. Berne: Peter Lang Verlag.
- Wagner, René-Louis, and Jacqueline Pinchon. 1973. *Grammaire du français classique et moderne*. 2<sup>e</sup> éd. revue et corrigée. Paris : Hachette.
- Wickham, Hadley. 2016. *ggplot2—Elegant Graphics for Data Analysis*. Springer. <https://www.springer.com/gp/book/9780387981413>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2021. *dplyr: A grammar of data manipulation (1.0.5)*. <https://CRAN.R-project.org/package=dplyr>.
- Willems, Dominique, and Filip Verroens. 2019. Les structures pronominales en français : Polyvalence syntaxique, ambiguïté sémantique et phénomènes de grammaticalisation. Les cas de *se mettre* et *se voir*. *Langages* 216: 103–119.
- Wilmet, Marc. 1998. *Grammaire critique du français*. 2<sup>e</sup> éd. Paris: Hachette.

## CORPORA

FT (Frantext): corpus of literary, philosophical, scientific and technical textual data  
LM (Le Monde): corpus made up of archives of the French newspaper *Le Monde*