Mapping Lexical Semantics onto Syntactic Structure:
The Problem of Unaccusative Mismatches in Romance Languages*

Géraldine Legendre¹ & Antonella Sorace²

¹Department of Cognitive Science, Johns Hopkins University
²Department of Theoretical and Applied Linguistics, University of Edinburgh
³legendre@cogsci.jhu.edu
⁴antonell@ling.ed.ac.uk

We propose that the interface between lexical-aspectual semantics and syntax is subject to violable mapping constraints as proposed in Optimality Theory (Prince & Smolensky [1993] 2004). This proposal is applied to an OT analysis of unaccusative mismatches in Romance languages in the domain of auxiliary selection in periphrastic past tenses.

1. Introduction

The problem of mapping lexical semantics onto syntactic structure became particularly acute with the Unaccusative Hypothesis (Perlmutter 1978) and its claim that intransitive verbs divide into two subsets with distinct syntactic properties. The single argument of unaccusative verbs is an underlying or deep direct object, and thus displays many syntactic properties of direct objects of transitive verbs; in contrast, the single argument of unergative verbs is a subject at all levels of representation, and thus displays the same syntactic behavior as the subject of transitive verbs. This syntactic difference is typically represented configurationally as in (1).

* The material in this article was presented by the first author at guest lectures at Seoul National University and the Korean Circle for Generative Grammar in November 2002. She thanks the audiences for their interest and comments and the organizers for their warm hospitality.

(1) Intransitive structures (Burzio 1986)

Unergative: \[ NP \ [v_p \ V] \]

Unaccusative: [\[ v_p \ V \ NP] \]

The simplicity and elegance of Perlmutter’s Unaccusative Hypothesis stands in sharp contrast with the many, largely unsuccessful, attempts at formulating a solid and systematic semantic basis for such a syntactic distinction and establishing its cross-linguistic validity.

The earliest formulations of the Unaccusative Hypothesis noted that the distinction is systematically related to certain semantic characteristics of the predicate: ‘agentivity’ tends to correlate with unergativity and ‘patienthood’ correlates with unaccusativity (Dowty 1991, Perlmutter 1978). Much subsequent research has shown, however, that the alignment between syntactic and semantic properties is not 100%; nor is it as consistent as originally predicted (Rosen 1984). Some verbs with similar semantics have different syntactic behavior across languages: for example, blush is unaccusative in Italian but unergative in French (and Dutch), on the basis of their auxiliary selection in periphrastic past tenses: Italian arrossire ‘blush’ selects esserelître (E) ‘be’ while its French counterpart rougir selects averelavoir (A) ‘have’. Some verbs are classified as both unaccusative and unergative by the same diagnostic: for example, Italian continuare ‘continue’ and French paraître ‘appear’ can take both auxiliary E and A. Within a given language syntactic tests do not overlap completely either. This is especially true of French where auxiliary selection identifies only a small subset of unaccusative verbs, compared with participial constructions — the syntactic test which best identifies the subclass of unaccusatives in French (Legendre 1989; Legendre & Sorace in press):

(2) F a. La neige a fondu pendant la nuit. (auxiliary: A)
‘The snow melted during the night’

F b. La neige fondu, toutes les stations de ski ont fermé.
‘The snow melted, all the ski resorts closed down’

Context too may play a role. It is well-known that correre ‘run’ selects E or A in Italian, depending on the presence of a stated goal. Yet the role of context is
not systematic across languages. French *courir* ‘run’ selects A, regardless of whether a goal is specified or not.

(3) I
   a. Il bambino è corso a scuola.  
   F b. L’enfant a couru à l’école.  
   ‘The child ran to school’

(4) I
   a. Il bambino ha corso nel giardino.  
   F b. L’enfant a couru dans le jardin.  
   ‘The child ran in the garden’

Nevertheless, a substantial body of research has shown that these ‘unaccusative mismatches’ are problematic only to the extent that one expects unaccusative and unergative verbs to represent syntactically and semantically homogeneous classes. Most of the syntactic diagnostics of unaccusativity/unergativity (e.g. auxiliary selection in Italian, impersonal passives in Dutch, resultative constructions in English) do tend to identify semantically coherent subsets of verbs (Levin & Rappaport Hovav 1995).

It is the formal nature of this tendency that we explore in this paper. Specifically, we argue that we can make genuine headway in understanding the complex facts of French in the context of Romance variation if we adopt the optimality-theoretic premise that well-formedness constraints on the mapping between the lexicon and syntax are universal but soft and highly conflicting. For example, verbs denoting existence of state select different perfect auxiliaries in the two languages: E in Italian vs. A in French. In our terms, such variation results from re-ranking a single constraint with respect to all others in the universal constraint hierarchy defining (part of) UG.

The paper is organized as follows. Section 2 reviews existing solutions to the lexicon/syntax mapping problem, including the Auxiliary Selection Hierarchy (Sorace 2000) which forms the basis for the OT analysis developed in Section 3. Section 4 considers general predictions of the proposed analysis plus a specific case: auxiliary selection in the history of Spanish culminating in the disappearance of *be*. Section 5 sketches out a extension to the OT analysis to handle the problem of indeterminacy of auxiliary choice for some verb classes in French and Italian. Section 6 offers concluding remarks.
2. Solutions to the Mapping Problem

In very general terms 25 years of research on the semantic basis of the unergative/unaccusative distinction have revealed the primacy of lexico-semantic and aspeccual features and emphasized the central role placed by telicicy in capturing regular patterns both across lexico-semantic verbs classes and across languages. What remains are controversies about the significance of other features or feature bundles, the theoretical status of verb classes, as well as the formal apparatus necessary to provide an explicit typology of possible lexicon/syntax mappings. Several classes of approach to the mapping problem have emerged in the last decade or so which we examine in turn. Because Sorace’s Auxiliary Selection Hierarchy was designed to solve the Romance problem it is discussed in much greater details than its competitors.

2.1. The Projectionist Approach

Levin & Rappaport Hovav are the leaders of what has become known as the ‘projectionist’ approach (see Sorace, in press, for discussion). They maintain that the lexical semantics of a verb deterministically specifies the hierarchical classification of its arguments, and that this in turn produces the syntactic behavior associated with unaccusativity or unergativity (Hale & Keyser 1986; Levin & Rappaport Hovav 1992, 1994, 1995, in press, among others).

The most comprehensive account of this type is Levin & Rappaport Hovav’s (1995) model based on English, in which a set of four linking rules map lexical semantic components of verb meaning onto positions at argument structure.

(5) Linking rules (Levin & Rappaport Hovav 1995)

i. Immediate Cause linking Rule: The argument of a verb that denotes the immediate cause of the eventuality described by that verb is its external argument (S).

ii. Directed Change Linking Rule: The argument of a verb that corresponds to the entity undergoing the directed change described by that verb is its direct internal argument (O).

iii. Existence Linking Rule: internal argument (O).

iv. Default Linking Rule: internal argument (O).
Within this approach, verbs with variable behavior have different meanings, and therefore different lexical semantic representations, each with its own regular argument structure realization.

Confronted with the complexities of Romance auxiliary selection, the projectionist approach faces at least two challenges — accounting for cross-linguistic variation without resorting to systematic duplication in the lexicon and accounting for French for which it overwhelmingly makes incorrect predictions. In particular, Rule ii incorrectly predicts that verbs of change of state (externally caused brûler ‘burn’, geler ‘freeze’, fondre ‘melt’, pourir ‘rot’, etc. and internally caused rougir ‘blush’, pâlir ‘become pale’, etc.) select E. Rule iii incorrectly predicts that verbs of existence (exister ‘exist’, durer ‘last’, être ‘be’, etc.) select E. Rule iv incorrectly predicts that non-agentive manner of motion verbs (rouler ‘roll’, rebondir ‘bounce’, tournoyer ‘whirl’, etc.) select E.

2.2. The Constructional Approach

Alternatives to the projectionist view have gained ground in recent years. Collectively they can be identified as ‘constructional’ approaches (Arad 1998; Borger 1994, 1998; Cummins 1996; van Hout 1996, 2000; McClure 1995; etc.). These models regard unaccusativity and unergativity not as lexical properties of verbs, but rather as clusters of properties derived from the syntactic configurations in which verbs appear, which in turn determine their aspectual interpretation. Since the lexical entry of verbs does not contain any specification of whether an argument is internal or external, any verb is free to enter into more than one syntactic configuration and consequently to receive multiple aspectual interpretations.

Unlike the projectionist model, the constructional approach predicts flexibility in the syntactic realization of arguments, but at the price of overgeneration. Constraints on overgeneration therefore have to be present at other levels (e.g. Cummins 1996, van Hout 1996). The constructional model is also a direct challenge to the Universal Alignment Hypothesis (UAH, Perlmutter 1978) and the Uniformity of Theta Assignment Hypothesis (UTAH, Baker 1988) according to which the mapping between thematic relations (agent, patient, etc.) and underlying syntactic configuration is invariable and universal.
In the specific context of Romance auxiliary selection the constructional approach incorrectly predicts that any verb may select E or A, depending on context. The prediction is incorrect: some verb classes are highly variable cross-linguistically while others are not. In all languages which allow two auxiliaries verbs of change of location (come, arrive, etc.) select E, not A while verbs of volitional activities (work, yell, etc.) select A, not E. Variation exists among other classes of verbs. This uneven variability is discussed in detail next.

2.3. The Auxiliary Selection Hierarchy (Sorace 2000)

Among the main Romance languages, French and Italian are the only languages which still make use of two auxiliaries in forming compound tenses.\(^1\) Most other languages have dropped their counterparts of E altogether, replacing it with a counterpart of A or some other auxiliary at some point in their history (e.g. Spanish haber, Catalan (Barcelona) haver, Portuguese tener). Romanian does make use of two auxiliaries but they do not alternate as markers of one and the same tense.

The starting point of Sorace’s Auxiliary Selection Hierarchy (ASH) is a set of facts which characterize auxiliary selection in Romance and Germanic languages: (a) across languages, some verbs tend to show consistent unaccusative/unergative behavior, whereas others do not; (b) within languages, some verbs are invariably unaccusative/unergative regardless of context, whereas others exhibit variation.

Sorace (et al.)’s studies provide supporting evidence for these generalizations, mostly based on experiments testing native speakers’ intuitions about auxiliary selection (perhaps the best known diagnostic of unaccusativity) in various languages that have a choice of perfective auxiliaries (such as Dutch, German, Italian, and Paduan). In all these languages — and to some extent in French, unaccusative verbs tend to select the counterpart of E and unergative verbs tend to select the counterpart of A. However, native intuitions on auxiliaries are categorical and consistent for certain types of verb, but much

\(^1\)Among less studied Romance languages Occitan, Piedmontese, Sardinian, and Catalan spoken outside of Barcelona maintain two auxiliaries.
less determinate for other types. For example, native speakers have a very strong preference for counterparts of E with change of location verbs, but express a weaker preference for the same auxiliary (or have no preference at all) with stative verbs.

Sorace’s 2000 account of these systematic differences within the syntactic classes of unaccusative and unergative verbs is that there exists a hierarchy which distinguishes ‘core’ unaccusative and unergative monadic verbs from progressively more ‘peripheral’ verbs. This hierarchy, which is based on (potentially universal) aspectual parameters, places the notion of telic dynamic change at the core of unaccusativity and that of agentive non-motional activity at the core of unergativity. The extremes of the hierarchy thus consist of maximally distinct core verbs - verbs of change of location (e.g. arrivare/arriver ‘arrive’) and verbs of agentive non-motional activity (e.g. lavorare/travailler ‘work’) - which consistently display the greatest degree of consistency in auxiliary selection. In contrast, peripheral verb types between the extremes are susceptible to variation. The overall hierarchy is represented in (6).

(6) The Auxiliary Selection Hierarchy

CHANGE OF LOCATION Selects E: essere/être (least variation)
CHANGE OF STATE
CONTINUATION OF A PRE-EXISTING STATE
EXISTENCE OF STATE
UNCONTROLLED PROCESS
CONTROLLED PROCESSES (MOTIONAL)
CONTROLLED PROCESS Selects A: avere/avoir (least variation)
(NON MOTIONAL)

Verbs at the extremes of the hierarchy (‘core’ verbs) are change of location verbs at the E end and non-motional process verbs at the A end. They are characterized by the following properties:

- categorical/consistent syntactic behavior across languages
- consistent behavior within individual languages; insensitivity to compositional properties of the predicate
- determinacy of native speakers’ intuitions
- primacy in acquisition
• diachronic stability

Let us examine some evidence in support of these generalizations, focusing in particular on the first three (for a full discussion see Sorace 2000, in press).

2.3.1. Core Verbs

Core verbs tend to be categorical and consistent in auxiliary selection across languages/language varieties. This is exemplified in (7)-(8), which show that the auxiliary selected by change of location verbs in the present perfect is E, and that selected by non-motional process verbs is A, in all the languages that have a choice of auxiliaries. S stands for Sardinian.

(7) I a. Paolo è venuto / *ha venuto in ritardo.
   ‘Paul came late’
   F b. Ma sœur est arrivée / *a arrivé hier.
   ‘My sister arrived yesterday’
   S c. Maria est / *at arrivata a domo.
   ‘Mary arrived home’

(8) I a. I delegati hanno parlato / *sono parlati tutto il giorno.
   F b. Les délégues ont parlé / *sont parlés toute la nuit.
   S c. Los profesores ont faeddadu / *son faeddados totu su die.
   ‘The delegates spoke all night long’

Core verbs display consistent behavior within individual languages; in particular, they tend to select the same auxiliary regardless of the contribution of other aspectual or thematic elements in the sentence in which they appear. So in (9) arrivare ‘arrive’ selects E (sono/sont) even though the predicate is atelic; the verb cadere ‘fall’ in (10a) selects E despite the fact that the event described by the verb clearly denotes intentionality, just as it does when the event is clearly unintentional (10b). Similarly, the verb lavorare ‘work’ selects A (hannolont) regardless of the telicity of the predicate, as in (11). Similar remarks apply to their French counterparts.

(9) I a. Sono arrivate lamentele in continuazione.  atelic predicate
Mapping Lexical Semantics onto Syntactic Structure:

F b. Des plaintes sont arrivées continuellement.
‘Complaints continuously came in’

(10) I a. Maria è caduta apposta per non andare a lavorare. agentive
F b. Maria est tombée volontairement pour ne pas aller travailler.
‘Mary fell down on purpose so as not to go to work’

I c. Il vaso è caduto dal tavolo. non-agentive
F d. Le vase est tombé de la table.
‘The vase fell down from the table’

(11) I a. I poliziotti hanno lavorato fino all’alba. telic predicate
F b. Les policiers ont travaillé jusqu’à l’aube.
‘The police worked will dawn’

The data from studies on other languages (e.g. Paduan, Cennamo & Sorace 1999) confirm that, in general, inherent lexical aspect determines auxiliary choice with core verbs, whereas compositional aspect (i.e. the event structure of the whole predicate) affects auxiliary selection with peripheral verbs. These findings support the conclusion that auxiliary selection with core verb types is a lexical phenomenon and is relatively insensitive to compositional factors. The degree of sensitivity to these factors increases for non-core verb types as they get more distant from the core.

Native speakers of languages with auxiliary selection have clear and determinate intuitions on core verbs; they categorically accept sentences in which these verbs appear with the ‘correct’ auxiliary and reject those in which they appear with the ‘wrong’ auxiliary. Evidence of differential judgments is particularly strong for Italian (Sorace 1993a, 1993b, 1995a; Bard, Robertson and Sorace 1996 for experimental evidence). Furthermore, descriptive studies of Italian (e.g. Berruto 1987; Rohlfis 1969) indicate that there is more variation in auxiliary usage for peripheral verbs than for core verbs, which is consistent with the predictions of the hierarchy.

The ASH is further supported by developmental data. Auxiliary selection with core verbs is acquired early both in first and second language acquisition. Data from the acquisition of Italian as a non-native language show that the syntactic properties of auxiliary selection are acquired first with core verbs and
then are gradually extended to more peripheral verb types (Sorace 1993a, 1995a). Moreover, Italian learners of French find it more difficult to acquire *avoir* as the auxiliary for verbs closer to the core than for peripheral verbs (Sorace 1993b, 1995b), and do not completely overcome this difficulty even at the advanced level. These developmental regularities can be explained by assuming that the acquisition of the syntax of unaccusatives crucially depends on the internalization of two elements: one is the hierarchical ordering of meaning components, and the other is the lexicon-syntactic mapping system instantiated by the target language.

A cursory look at the early acquisition of French verbs by young Grégoire (Champaud Corpus, available from the CHILDES Database, McWhinney and Snow 1985) confirms the general findings. In his earliest 4 files (Age: 1; 9-1; 10) the only intransitive verbs Grégoire uses are unaccusative; he produces passé composé forms with the correct auxiliary (E) with verbs of location first (specifically *tomber* ‘fall’, *monter* ‘go up’, *partir* ‘leave’, in this order). The first unergative verbs to show up in the passé composé (A) are controlled motional processes *bouger* ‘move’ (2; 0; file # 5) and *rouler* ‘move/roll (for a car)’ (2; 3; file #7).

Finally, core verbs tend to be diachronically stable. There is evidence from studies on the historical development of auxiliaries in Romance (e.g. Benzing 1931, Tuttle 1986) showing that core verb types tend to be the last to be affected by the replacement of auxiliaries derived from Lat. esse (E) with those derived from *habere* (A) whereas peripheral verb types are the most vulnerable to the change (see further discussion in Section 4). A recent study by Cennamo (1999) suggests that the development of reflexives *selsibi* in Late Latin as markers of split intransitivity followed a path largely consistent with the ASH.

### 2.3.2 Intermediate (non-core) Verbs

While core verbs tend to be categorical in their auxiliary selection behaviour, non-core verbs show increasing variation. The greater flexibility of these verbs is illustrated here with Italian and French examples (for further cross-linguistic evidence see Sorace 2000).

A class that exhibits regular alternations is that of verbs denoting 'indefinite change' in a particular direction (e.g. *monter* ‘go up’), change of condition (e.g. *faner* ‘wilt’), appearance (e.g. *apparaître* ‘appear’). E is strongly preferred by
these verbs in Italian, but A (ha) is not completely rejected (as in (12c,e)). The strength of preferences is a function of the (+/-) inherent telicity of the verb: as the Italian sentences in (12)-(13) show, many of these verbs allow two readings, one telic and one atelic, which may be disambiguated by the context.

(12) I a. La popolarità del governo è scesa / ha sceso notevolmente.
F b. La popularité du gouvernement a (visiblement) monté.
   ‘The government popularity has visibly increased’
I c. Mia figlia è cresciuta / ?*ha cresciuto molto quest’anno.
F d. Ma fille a (beaucoup) grandi (cette année).
   ‘My daughter grew up a lot (this year)’
I e. Lo spettro è apparso / ?*ha apparso nel castello.
F f. Le fantôme est apparu (dans le château).
   ‘The ghost appeared in the castle’

(13) I a. La pianta è / ha fiorita due volte quest’anno.
F b. La plante a fleuri deux fois cette année.
   ‘The plant bloomed twice this year’
I c. I pomodori sono marciti /hanno marcito al sole.
F d. Les tomates ont pourri au soleil.
   ‘The tomatoes rotted in the sun’
I e. Il girasole è / ha finalmente germogliato.
F f. Le tournesol a enfin fleuri.
   ‘The sunflower finally bloomed’

---

2 The diacritics in the examples, here and throughout, refer to the degrees of unacceptability of a sentence in terms of strength of preference that native speakers have for one auxiliary over the other. The do not refer to the normative acceptability of sentences in terms of prescriptive grammars. So a sentence marked as "?*" may be ungrammatical according to a prescriptive grammar of Italian, but is judged by native speakers as more acceptable than a sentence marked with "*". For a discussion on the quantification of relative judgments of linguistic acceptability, and experimental results, see Bard, Robertson & Sorace (1996).
The gloss reveals that their French counterparts select A in the same contexts, with two exceptions: monter ‘go up’ (and descendre ‘go down’) select E or A for most speakers depending on the agentivity of its subject. For some speakers either auxiliary is possible in these contexts with a subtle change in meaning. E emphasizes the fact that the goal is reached; A conveys more of the difficulty during the ascension and somewhat unexpected success of reaching the summit.

(14) F a. Pierre est/a monté jusqu’au sommet.
   ‘Peter went up to the top’

   F b. La température a monté/est montée\(^3\) pendant la journée.
   ‘The temperature went up/increased during the day’

\textit{Apparaître} ‘appear’ typically selects E but it is not uncommon to see it with A (15a). In fact, the class of verbs of appearance is the class that displays the most variation in French. Yet, the change in auxiliary does not seem to correlate with a change in meaning for this class.

(15) F a. C’est ici que le petit prince a apparu sur terre. \quad (Saint-Exupéry)
   ‘It’s here that the little prince appeared on earth’

   F b. Ambrose Pierce a/est disparu en 1913. \quad (Cummins 1996: 39)
   ‘AP disappeared in 1913’

   F c. Le dernier livre de Chomsky a/est paru en 1995.
   ‘Chomsky’s last book appeared in 1995’

   F d. Eve a/est passé(e) de la chambre à coucher à la salle de bain.
   \quad (cited in Ruwet 1988)
   ‘Eva went from the bedroom to the bathroom’

Verbs denoting continuation of a pre-existing condition (e.g. rester ‘stay’) are less determinate in Italian: E is preferred but A is not ruled out

\[^3\] Use of E in French entails gender agreement on the past participle form of the lexical verb, typically spelled as \(e\) if feminine.
categorically, and is in fact accepted with many of these verbs. The agentivity of the subject correlates with the degree of acceptance of A (see the contrast in (16), suggesting that these verbs, unlike core verbs, are sensitive to the feature contributed at the predicate level.)

   F  b. Je suis resté seul une fois de plus.
       ‘I stayed alone once more’
   I  c. La discussione è / ?ha durato a lungo.  non agentive
   F  d. La discussion a duré pendant longtemps.
       ‘The discussion lasted for a long time’
   I  e. Il presidente è / ha durato in carica tre mesi.  agentive
   F  f. Le doyen a duré trois mois dans son poste.
       ‘The dean lasted three months in his job’
   I  g. I miei genitori sono / ?hanno sopravvissuto alla guerra.
   F  h. Mes parents ont survécu à la guerre.
       ‘My parents survived the war’
   I  i. Questa atteggiamento è / ?ha persistito per troppo tempo.
   F  j. Cette attitude a persisté pendant trop longtemps.
       ‘This attitude persisted for too long’
   I  k. Gianni *è / ha persistito nella sua ostinazione.
   F  l. Jean a persisté dans son obstination.
       ‘John persisted in his stubbornness’

The French counterparts to (16) invariably select A with one exception: rester ‘stay’. In fact, all remaining verb classes discussed below invariably select A, in sharp contrast to the variation displayed in Italian.

Stative verbs (including both verbs of physical and abstract existence and psychological verbs) are the most indeterminate in Italian, consistent with the findings from other studies. Auxiliary alternations (some restricted to regional or non-standard varieties) are shown in (17), (18).
(17) I a. I primi mammiferi sono esistiti / ??hanno esistito molti milioni di anni fa.
F b. Les premiers mammiferes ont existé il ya a des millions d’années.
‘The first mammals existed millions of years ago’

I c. Lo zucchero non è bastato / ??ha bastato per fare la torta.
F d. Le sucre n’a pas/a suffit pour faire le gateau.
‘The sugar wasn’t enough (lit. didn’t suffice) to make the cake’

I e. Il film è sembrato / ??ha sembrato troppo violento a tutti gli spettatori.
F f. Le film a semblé trop violent à tous les spectateurs.
‘The film seemed too violent to the entire audience’

(18) I a. Questo palazzo ha appartenuto / è appartenuto alla mia famiglia.
F b. Cet hôtel particulier a appartenu à ma famille.
‘This mansion belonged to my family’

I c. I viveri sono scareaggiati / hanno scareaggiato tra i terremotati.
F d. Les vivres ont beaucoup diminué parmi les victimes du tremblement de terre.
‘Provisions greatly diminished among the earthquake victims’

I e. Il partito è / ??ha sussistito senza i contributi dei politici.
F f. Le parti a subsisté sans contributions des politiciens.
‘The party subsisted with no contributions from politicians’

I g. La sua dichiarazione non è servita / ??ha servito a nulla
F h. Sa declaration n’a servi à rien.
‘His declaration served no purpose’

The use of A induces an agentive reading, whereas E does not. So in (19), the Italian verb *mancare* ‘miss’ is understood as intentional in (a,b) and non-intentional in (b,c). Its French counterpart selects A, regardless of intentionality.
(19) I a. Il soldato è mancato all’appello. \textit{non-agentive}
F b. Le soldat a manqué à l’appel.
   ‘The soldier was absent from (lit. missed) roll call’
I c. Il presidente ha mancato all’appuntamento. \textit{agentive}
F d. Le président a manqué au rendez-vous.
   ‘The president missed the appointment’

Peripheral verbs closer to the ‘unergative’ core include verbs denoting
motional processes (e.g. \textit{jump, swim}, etc.). Native Italian intuitions are less
determinate: A is preferred but E is not completely rejected, as shown in (20), (21).

(20) I a. I bambini hanno saltato / ?*sono saltati in giardino tutto il
   pomeriggio.
F b. Les enfants ont sauté dans le jardin tout l’après-midi.
   ‘The children jumped in the garden all afternoon long’

   F b. Michèle a couru plus vite que n’importe qui.
   ‘Michelle ran faster than anyone else’
   I c. Paola ha nuotato / ?*è nuotata fino all’altra sponda.
   F d. Paola a nagé jusqu’à l’autre rive.
   ‘Paola swam to the other bank’

The effect of agentivity/volitionality on auxiliary selection in Italian (but not
in French) is shown in (22), where A is the preferred auxiliary with a human
subject, E is the preferred one with an inanimate subject.

(22) I a. Il pilota ha / ?è atterrato sulla pista di emergenza.
   F b. Le pilote a attéré sur la piste d’urgence.
   ‘The pilot landed on the emergency runway’
   I c. L’elicottero è / ?ha atterrato sul tetto del grattacielo
   F d. L’hélicoptère a attéri sur le toit du gratte-ciel.
   ‘The helicopter landed on the skyscraper roof’
Next, the hierarchy includes various types of uncontrolled processes (such as bodily functions (e.g. suer ‘sweat’), involuntary reaction (e.g. trembler ‘tremble’) and emission (e.g. cliqueter ‘rattle’). (for definitions of controlled vs. uncontrolled processes, see Levin & Rappaport Hovav 1995). These verbs are internally caused but tend to be non-volitional.

(23) I a. Il convincimento politico ha tentennato / ? è tentennato anche nei più anziani.
F b. Les convictions politiques ont vacillé même chez les personnes les plus âgées.
‘Political convictions wavered even among older people’
I c. Paolo ha tentennato / *è tentennato a lungo prima di prendere una decisione
F d. Paul a vacillé pendant longtemps avant de prendre une décision.
‘Paul wavered for a long time before making a decision’
I e. La terra ha tremato / ?è tremata.
F f. La terre a tremblé.
‘The earth shook (lit. trembled)’
I g. Mario ha tremato / *?è tremato dallo spavento.
F h. Mario a tremblé de peur.
‘Mario shook (lit. trembled) with fear’
I i. Il mendicante ha rabbrividito / è rabbrividito dal freddo.
F j. Le mendiant a tremblé de froid.
‘The beggar shivered (lit. trembled) with cold’

(24) I a. L’innesto non è attecchito / ha attecchito.
F b. La greffe n’a pas pris.
‘The transplant didn’t take’
I c. L’acqua ha / ?è scarceggiata.
F d. L’eau s’est faite rare.4

4This predicate is morphologically reflexive in French, which necessitates the use of
‘Water became rare’

I e. La bicicletta ha / ?è sbandata senza preavviso.
F f. La bicyclette a soudain dérapé.
‘Suddenly the bicycle skidded’

(25) I a. La sveglia ha / ?è squillata.
F b. Le réveille-matin a sonné.
‘The alarm clock rang’

I c. L’eco ha / è risuonato.
F d. L’écho a résonné.
‘The echo resonated’

I e. Il tuono ha / è rimbombato.
F f. Le tonnerre a grondé.
‘The thunder rumbled’

To sum up, auxiliary selection in Italian displays a gradient sensitivity to the aspeccual and lexical-semantic properties of individual verbs, which is uniquely captured by the ASH. Telicity is the main factor that separates verbs which select E from verbs which select A. Agentivity is a secondary factor that further differentiates among verbs selecting A. Core verbs at the extremes of the hierarchy tend to select the same auxiliary categorically both within and across languages; verbs between the extremes are less specified with respect to telicity and agentivity, and it is among these verbs that most cases of ‘unaccusative mismatches’ are found.

French too displays some gradient sensitivity to the aspeccual and lexical-semantic properties of individual verbs, with one significant difference. The core unaccusative verbs are a small subset of their Italian counterparts and the resulting cut-off point between intransitive verbs selecting E and those selecting A is much higher in the hierarchy, as shown in Table 1 (* indicates variation as discussed above).

E. See Legendre & Sorace (in press) for further discussion and a straightforward extension of the present analysis.
<table>
<thead>
<tr>
<th>French</th>
<th>Italian</th>
<th>Verb classes and sample verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>E</td>
<td>Change of location: <em>arrive, come</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change of state</td>
</tr>
<tr>
<td>E</td>
<td>E</td>
<td>a. Change of condition: <em>die</em></td>
</tr>
<tr>
<td>E*</td>
<td>A</td>
<td>b. Appearance: <em>(dis)appear</em></td>
</tr>
<tr>
<td>E*</td>
<td>E</td>
<td>c. Indefinite change in a particular direction: <em>go up, go down</em></td>
</tr>
<tr>
<td>A</td>
<td>E</td>
<td><em>wilt, rot</em></td>
</tr>
<tr>
<td>A</td>
<td>E*</td>
<td><em>worsen</em></td>
</tr>
<tr>
<td>A</td>
<td>E*</td>
<td>States</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a. Continuation of pre-existing state: <em>last</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Existence of state:</td>
</tr>
<tr>
<td>A</td>
<td>E</td>
<td><em>be</em></td>
</tr>
<tr>
<td>A</td>
<td>E*</td>
<td><em>exist, suffice</em></td>
</tr>
<tr>
<td>A</td>
<td>A*</td>
<td>Uncontrolled processes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a. Emission: <em>resound</em></td>
</tr>
<tr>
<td>A</td>
<td>A*</td>
<td>b. Involuntary actions: <em>shiver</em></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>c. Bodily functions: <em>sweat</em></td>
</tr>
<tr>
<td>A</td>
<td>A*</td>
<td>Controlled processes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a. Motional: <em>swim</em></td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>b. Non-motional: <em>work, yell</em></td>
</tr>
</tbody>
</table>

The high cut-off point in French entails that variation and instability occurs closer to the top, i.e. among change of state verbs. As many verb classes from the bottom up have already switched to A, there is less variation and gradience than in Italian. Viewed from the perspective of the history of Romance languages, French appears to be relatively close to eliminating E as an alternating auxiliary.

The ASH challenges existing theories of the syntax-lexicon interface. It cannot be accommodated within a projectionist account because it would entail too much duplication in the lexicon, and it does not fit a constructional account either because the amount of variation is related to specific verb types. At the same time, it has features of both accounts: like the projectionist approach, it
assumes a systematic relation between the syntax of auxiliary selection and the semantics of individual verbs; like the constructional approach, it allows for verbs (though not all) to have multiple syntactic projections.

3. Optimizing Auxiliary Selection

Establishing that cross-linguistic variation in auxiliary selection is best understood in terms of a hierarchy of lexico-semantic verb classes still leaves important theoretical questions unanswered. The ASH is a generalization which reveals different cut-off points for the unergative/unaccusative distinction cross-linguistically (as revealed in auxiliary selection in French and Italian). It does not automatically translate into a set of mapping rules referring to the verb classes in Table 1, for two main reasons. A single verb class may not map onto a single auxiliary — as is the case for change of state verbs in French: *die* selects E, *rot* selects A. More important still is the fact that these classes do not by themselves reveal what is common to two verb classes selecting one and the same auxiliary.

We propose that the ASH arises from an optimization-based view of grammar whereby the verb classes listed in Table 1 and the hierarchy itself *emerge* from a competition among soft constraints on mapping a given lexico-semantic or aspectual feature (e.g., *telicity, direction*, etc.) onto a syntactic configuration. The basic form of the constraints given in (26) highlights the independent issue of instantiating unergative and unaccusative configurations in particular theories of syntax.

(26) a. *S/X: “don’t map feature X onto an unergative configuration”

b. *O/Y “don’t map feature Y onto an unaccusative configuration”

3.1. Event Features

The lexico-semantic and aspectual properties widely implicated in unaccusativity phenomena are telicity, directed change, change of state, motion, displacement (e.g. Levin & Rappaport 1995, Zaenen 1993, Dowty 1991, Van Valin 1990, Sorace 2000, etc.) and homogeneity (McClure 1995). (27) represents the smallest set of features necessary to exhaustively characterize auxiliary selection in Romance. Further discussion of these
features follows summarizing Table 2.

(27) Predicate features: INHERENT TELICITY, DIRECTION, STATE, INHERENT DISPLACEMENT, INTERNAL MOTION, HOMOGENEITY, INHERENT VOLITIONALITY

We proceed with some selective discussion of our choice of features with the understanding that the relevance of most features has a solid basis based on the relevant literature cited in this paper. Among those is inherent telicity determined on the basis of diagnostic tests like occurrence with the adverbial phrase for an hour/ an hour (Dowty 1979). On the basis of these tests change of location verbs like arriver ‘arrive’ are telic while controlled processes like travailler ‘work’ are atelic.

(28) F a. Pierre est arrivé chez lui en 1 h.
   ‘Peter arrived home in one hour’

F b. Pierre a pris/mis 1 h pour arriver chez lui.
   ‘Peter took one hour to arrive home’

F c. *Pierre est arrivé pendant 1 h.
   ‘Peter arrived for one hour’

Table 2. Featural composition of French and Italian monadic verbs

<table>
<thead>
<tr>
<th>Aux</th>
<th>Features of predicate →</th>
<th>DIS</th>
<th>-HOM</th>
<th>TEDIR</th>
<th>ST</th>
<th>-VO</th>
<th>MO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Change of location: arrive</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Change of state</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a. Change of condition: die</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>E*</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. Appearance: (dis)appear</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>c. Indefinite change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>in a particular direction:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E*</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>go up</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>A</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rot</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>A</td>
<td>E*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>worsen</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Mapping Lexical Semantics onto Syntactic Structure:

<table>
<thead>
<tr>
<th>States</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Continuation of a pre-existing state: last</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>A E*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Existence of state: exist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>A E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncontrolled processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A A*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>a. Involuntary actions: shiver</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A A*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>b. Emission: resound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>c. Bodily functions: sweat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Controlled processes                                                  |   |   |   |   |   |   |   |
| A A*                                                                   |   |   |   |   | - |   |   |
| a. Motional: swim                                                     |   |   |   |   |   |   |   |
| A A                                                                    |   |   |   |   | - |   |   |
| b. Non-motional: yell                                                 |   |   |   |   |   |   |   |

* = Both auxiliaries are possible (see Section 5 for further discussion)

(29) F a.  *Pierre a travaillé en 3 h.
  ‘Peter worked in three hours’

  F b.  *Pierre a pris/mis 3 h pour travailler.
  ‘Peter took three hours to work’

  F c.  Pierre a travaillé pendant 3h.
  ‘Peter worked for three hours’

A less traveled feature is that of homogeneity or the extent to which each subinterval of an event is identical (McClure 1995). Events as different as controlled activities travailler ‘work’ and change of state pourir ‘rot’ are homogenous in the following way. Each sub-event of an on-going event of working or rotting entails that the referent has worked or rotted at least a bit. In contrast, an on-going event of dying (because it canonically takes time to die) is not made of homogeneous sub-events: at any point of getting closer to the endpoint one cannot say that the agent has died (even a bit).

(30) Event homogeneity:
  F a.  travailler ‘work’
  Jean est en train de travailler.  ⇒  il a travaillé
‘John is (in the process of) working’    ‘he has worked’

F b. *pourir* ‘rot’
    Ta pomme est en train de pourir.     \Rightarrow    elle a pouri un peu
    ‘Your apple is rotting’             ‘it has rotted a bit’

F c. *mourir* ‘die’
    Jean est en train de mourir.       \Rightarrow    il est mort (un peu)
    ‘John is dying’                    ‘he has died (a bit)’

We need to distinguish change of state verbs like *die* from *rot* by at least one feature value because they select different auxiliaries in French. In the absence of the feature *homogeneity*, the featural profile of both verbs is identical, as shown in Table 2. We propose that the difference in auxiliary choice within this traditional verb class is due to the fact that *die* is non-homogeneous while *rot* is. Processes of rotting, dying, appearing or disappearing are all gradual processes. What distinguishes them is whether the gradual process is incrementally homogeneous or not.

Another important distinction to be made is between verbs which connote inherent displacement from point A to point B (e.g. aller à ‘go to’ and more generally change of location verbs) and verbs which connote internal motion (e.g. *nager* ‘swim’, *courir* ‘run’). Note that these types of motion can occur without displacement as revealed in the common expression *nager, courir sur place* ‘swim, run in place’. In Table 2 the feature *motion* distinguishes among controlled (volitional) processes e.g. *swim* vs. *yell*.

Finally, *inherent volitionality* distinguishes uncontrolled from controlled processes at the bottom of the hierarchy. Change of location and condition verbs (e.g. *venir* ‘come’, *mourir* ‘die’) are not inherently volitional. Volitionality, when present, is a property of their argument.

The remaining features in Table 2 are *direction* and *state*. The former distinguishes change of location and state verbs from the rest because an important component of their meaning is ‘directed change’ and the latter characterizes verbs whose meaning includes being in or reaching a state (incl. location at point B).

Using these features and their appropriate values we obtain a set-inclusion hierarchy, as represented in Table 2. What is crucial here is that verb classes
that select a different auxiliary be distinguished by at least one feature value and that these feature values express implicational relations (+DIS implies −HOM; −HOM implies +TE; +TE implies +DIR, etc.) This is true of the first six features, counting from the left end side. This distribution enables us to propose an OT analysis grounded in a Power Hierarchy (cf. Section 3.2) whose universal scope does not rely on any further stipulation.

### 3.2. Harmonic Alignment

One important outcome of much typological-functional research is the recognized existence of markedness relations (e.g. Croft 1990, Jakobson [1965], 1995, Keenan & Comrie 1977, Perlmutter 1983, Silverstein 1976, etc.) which express favored associations in languages of the world. A well-known one is grounded in the relational scale stated in (31).

(31) GF scale: subject (S) > direct object (O)

On the model of animacy (Local person > Pronoun 3rd > Human 3rd > etc.) and thematic (Agent > Patient) scales that get associated with (31) to express markedness relations (e.g. Aissen 2001), it is easy to formulate event scales for the features relevant to the unaccusative/unergative distinction. In fact, the telicity scale (atelic > telic) is assumed in Grimshaw (1990).

(32) Event feature scales:

<table>
<thead>
<tr>
<th>Displacement:</th>
<th>−DIS</th>
<th>&gt;</th>
<th>+DIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneity:</td>
<td>+HOM</td>
<td>&gt;</td>
<td>−HOM</td>
</tr>
<tr>
<td>Telicity:</td>
<td>−TE</td>
<td>&gt;</td>
<td>+TE</td>
</tr>
<tr>
<td>Directed change:</td>
<td>−DIR</td>
<td>&gt;</td>
<td>+DIR</td>
</tr>
<tr>
<td>State:</td>
<td>(ST)</td>
<td>&gt;</td>
<td>+ST</td>
</tr>
<tr>
<td>Inherent volitionality:</td>
<td>+VO</td>
<td>&gt;</td>
<td>−VO</td>
</tr>
<tr>
<td>Internal motion:</td>
<td>−MO</td>
<td>&gt;</td>
<td>+MO</td>
</tr>
</tbody>
</table>

By aligning two scales at a time we come up with a set of relations which express how marked the mapping of a certain feature — say TELIC — is with a certain grammatical function — say OBJECT. Such harmonic alignments (as they are known in OT) embody preferences for mappings between certain
properties across scales, i.e. markedness relations (Prince and Smolensky [1993] 2004; Aissen 2001). Note the change of symbol from ‘>’ to ‘ ’ in the sample harmonic alignments in (33).

(33) Harmonic alignments:
   a. O/telic > S/telic
   b. S/atelic > O/atelic
   c. etc.

   The mapping of [+telic] onto an O configuration is less marked than the mapping of [+telic] onto a S configuration

   Such alignments turn into a hierarchy of constraints or filters once their polarity is reversed (note again the change in symbol).

(34) Constraint alignments:
   a. *S/telic \succ *O/telic
      don’t map [+telic] onto a S configuration
      outranks (\succ)
      ‘don’t map [+telic] onto an O configuration
   b. *O/atelic \succ *S/atelic
   c. etc.

Putting all mapping constraints pertaining for example to a Subject configuration together we obtain a hierarchy or ranking of constraints which is hypothesized to be universally fixed. Such ranking is called a Power Hierarchy in the OT literature (Smolensky 1995, Legendre et al. 1998). The relevant technicalities are outlined in Section 3.3.

3.3. Formulating a *S Power Hierarchy

Consider again Table 2. Down to feature [(VOLITIONAL] Table 2 expresses implicational relations among feature values. We can thus define a set \( \mathcal{C} \) of “object-preferring” feature values as in (35) and state an encapsulated constraint \(*S/\mathcal{C} \) which is violated whenever a constraint in \{*S/f: f in \( \mathcal{C} \)\} is violated.

(35) a. \( \mathcal{C} \) (“object-preferring” feature values)
Mapping Lexical Semantics onto Syntactic Structure:

= \{+\text{DIS}, -\text{HOM}, +\text{TE}, +\text{DIR}, +\text{ST}, -\text{VO}\}

b. \(F = *S/\emptyset\): “Don’t assign S to an event with a \(\emptyset\)-feature”

The fact that a candidate violating \(*S/\emptyset\) six times is more marked than one violating it twice is implemented via standard OT \textit{Power Hierarchy} (Smolensky 1995; Legendre et al. 1998). It is expressed in its encapsulated fashion in \((36a)\) and in its non-encapsulated fashion in \((36b)\) where constraint names are based on the implicational relations in Table 2.

\((36)\) Universal mapping constraint hierarchy

\begin{align*}
\text{a. } & F^6 \gg \cdots \gg F^1 \quad \text{‘the Mapping Hierarchy’} \\
& \text{(event semantics/syntax mapping)} \\
\text{or} \\
\text{b. } & *S/+\text{DIS} \gg *S/-\text{HOM} \gg *S/+\text{TE} \gg \\
& *S/+\text{DIR} \gg *S/+\text{ST} \gg *S/-\text{VO}
\end{align*}

As shown in Table 3, it is more marked for \textit{arriver} ‘arrive’ to be assigned a S configuration than for \textit{suer} ‘sweat’ because \textit{arriver} ‘arrive’ violates constraint F 6 times while \textit{suer} violates it only once. Hence it is more likely that \textit{arriver} will be assigned a O configuration than \textit{suer}.

\begin{table}[h]
\centering
\caption{Markedness as determined by \(*S\) Power Hierarchy}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\textbf{Power Hierarchy} & \(*S/\emptyset^6\) & \(*S/\emptyset^5\) & \(*S/\emptyset^4\) & \(*S/\emptyset^3\) & \(*S/\emptyset^2\) & \(*S/\emptyset\) \\
\hline
\textbf{Constraint ranking} & \text{\textit{Slarriver} ‘arrive’} & \text{\textit{Slarriver} ‘arrive’} & \text{\textit{Slarriver} ‘arrive’} & \text{\textit{Slarriver} ‘arrive’} & \text{\textit{Slarriver} ‘arrive’} & \text{\textit{Slarriver} ‘arrive’} \\
\hline
\text{+DIS, -HOM, +TE, +DIR, +ST, -VO} & * & * & * & * & * & * \\
\hline
\text{\textit{Slsuer} ‘sweat’} & \text{\textit{Slsuer} ‘sweat’} & \text{\textit{Slsuer} ‘sweat’} & \text{\textit{Slsuer} ‘sweat’} & \text{\textit{Slsuer} ‘sweat’} & \text{\textit{Slsuer} ‘sweat’} & \text{\textit{Slsuer} ‘sweat’} \\
\hline
\text{–DIS, +HOM, –TE, –DIR, –ST, –VO} & & & & & & * \\
\hline
\text{\textit{Snager} ‘swim’} & \text{\textit{Snager} ‘swim’} & \text{\textit{Snager} ‘swim’} & \text{\textit{Snager} ‘swim’} & \text{\textit{Snager} ‘swim’} & \text{\textit{Snager} ‘swim’} & \text{\textit{Snager} ‘swim’} \\
\hline
\text{–DIS, +HOM, –TE, –DIR, –ST, +VO} & & & & & & \\
\hline
\end{tabular}
\end{table}
Recall that an additional feature is needed to distinguish motional from non-motional controlled processes: *nager ‘swim’ vs. *travailler ‘work’. The feature proposed is MO (INTERNAL MOTION). Because all verb classes above controlled motional processes have the value —MO and so does the lowest class on the ASH that includes *travailler, the feature MO does not stand in an implicational relation with all other features. The constraint *S/MO is therefore not part of the Power Hierarchy proper. The model is a hybrid one.

However, the constraint *S/MO has to be ranked below the lowest constraint in the Power Hierarchy to express the relative markedness of assigning a O configuration to the *nager vs. *travailler classes. Anticipating the diachronic evidence of auxiliary change in Spanish presented in Section 4, it is clear that *travailler is more unergative than *nager.\(^5\) *Nager violates *S/MO, *travailler violates no *S constraint.

Putting together the *S Power Hierarchy and *S/NO we obtain the *S constraint hypothesized to be universal.

\[(37)\] Universal *S constraint ranking:

\[\text{*S/+DIS} \gg \text{*S/-HOM} \gg \text{*S/+TE} \gg \text{*S/+DIR} \gg \text{*S/+ST} \gg \text{*S/-VO} \gg \text{*S/NO}\]

Obviously, if no constraint on mapping onto an O configuration ever entered the picture, an O configuration would always be optimal, we would have no cross-linguistic mismatches and no unergative verbs. While the harmonic alignments stated in (33) technically result in a hierarchy of *O constraints parallel to that of *S constraints we need only use the encapsulated version given in (38) for the purpose of explaining auxiliary selection in Romance.

\[(38)\] *O: “don’t map onto an unaccusative configuration”

### 3.4. Establishing the language-particular ranking of *O

It should be clear by now that the cross-linguistic choice of auxiliary for a given featural profile will be determined by the relative ranking of *O, i.e.

\(^5\) See also the data on $swim$ in Italian in Section 2.3.2. Selecting E is not completely rejected by native speakers.
where *O is interposed on the *S hierarchy. Thus our analysis rests on the claim that while the ranking of *S constraints is universally fixed that of *O is language-specific.

We take the input to optimization to be the featural description of individual predicates or predicate subclasses, as specified in Table 2. The candidate set simply consists of two candidates corresponding to the two auxiliaries, E and A. Note that a direct mapping from GF to selected auxiliary is assumed: S (A), O (E). The candidate auxiliary which wins in each optimization is the one which incurs minimal violations of the lower-ranked constraints.

A sample of tableaux is provided next which highlights the crucial optimizations responsible for the language-particular rankings of *O stated in (39).

**Tableau 1. pourir ‘rot’**

<table>
<thead>
<tr>
<th>Input: –DIS, +HOM, +TE, +DIR, +ST, –VO, –MO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. O (E)</td>
</tr>
<tr>
<td>Ab. S (A)</td>
</tr>
</tbody>
</table>

Despite the fact that it is telic, change of state verb *pourir* selects A. This motivates the sub-ranking *O ≫ *S/TE. But *mourir* selects E, despite the fact that it is also a telic change of state verb. This is where the constraint on mapping the feature HOMOGENEITY comes in: *S/–HOM ≫ *O, as shown in Tableau 2.

**Tableau 2. mourir ‘die’**

<table>
<thead>
<tr>
<th>Input: –DIS, –HOM, +TE, +DIR, +ST, –VO, –MO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. O (E)</td>
</tr>
<tr>
<td>b. S (A)</td>
</tr>
</tbody>
</table>
In Italian verbs of state select E. This entails that *O must be outranked by *S/ST. See Tableau 3. In contrast, [−ST] verbs select A. Thus *O ≫ *S/−VO. See Tableau 4.

**Tableau 3. esistere ‘exist’**

<table>
<thead>
<tr>
<th>Input: −DIS, +HOM, −TE, −DIR, +ST, −VO, −MO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
</tr>
<tr>
<td>*S/ DIS</td>
</tr>
<tr>
<td>Δa. O (E)</td>
</tr>
<tr>
<td><img src="image0" alt="Image" /></td>
</tr>
<tr>
<td><img src="image8" alt="Image" /></td>
</tr>
<tr>
<td><img src="image16" alt="Image" /></td>
</tr>
</tbody>
</table>

**Tableau 4. sudare ‘sweat’**

<table>
<thead>
<tr>
<th>Input: −DIS, +HOM, −TE, −DIR, −ST, −VO, −MO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
</tr>
<tr>
<td>*S/ DIS</td>
</tr>
<tr>
<td>a. O (E)</td>
</tr>
<tr>
<td><img src="image24" alt="Image" /></td>
</tr>
<tr>
<td><img src="image32" alt="Image" /></td>
</tr>
<tr>
<td><img src="image40" alt="Image" /></td>
</tr>
</tbody>
</table>

b. S (A)                                    |
| ![Image](image48) | ![Image](image49) | ![Image](image50) | ![Image](image51) | ![Image](image52) | ![Image](image53) | ![Image](image54) | ![Image](image55) |
| ![Image](image56) | ![Image](image57) | ![Image](image58) | ![Image](image59) | ![Image](image60) | ![Image](image61) | ![Image](image62) | ![Image](image63) |
| ![Image](image64) | ![Image](image65) | ![Image](image66) | ![Image](image67) | ![Image](image68) | ![Image](image69) | ![Image](image70) | ![Image](image71) |

These crucial cases result in the language-particular constraint rankings stated in (39).

(39) a. French:

```
*O

*S/DIS ≫ *S/(HOM ≫ *S/+TE ≫ *S/+DIR ≫ *S/+ST ≫ *S/−VO ≫ *S/MO
```

b. Italian:

```
*O
```

### 3.5. Verifying Predictions in French and Italian

For Italian the position of *O in the *S hierarchy entails that verbs that have any of the feature values [−DIS, −HOM, +TE, +DIR] select E. For example, peggiorare ‘worse’ selects E in Italian, as expected; see Tableau 5. Its French counterpart, however, is predicted to select A because *O outranks *S/DIR in
French; see Tableau 6.

**Tableau 5.** *peggiorare* ‘worsen’

<table>
<thead>
<tr>
<th>Italian</th>
<th>*S/ DIS</th>
<th>*S/ -HOM</th>
<th>*S/ TE</th>
<th>*S/ DIR</th>
<th>*S/ ST</th>
<th>*O</th>
<th>*S/ -VO</th>
<th>*S/ MO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aa. O (E)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. S (A)</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tableau 6.** *emprer* ‘worsen’

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. O (E)</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ab. S (A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

In both languages, controlled processes (motional and non-motional) are predicted to select A. That is because only *O and *S/MO are activated in the case of motional processes and only *O in the case of non-motional processes (e.g. *travailler/lavorare* ‘work’). Given that *O ≫ *S/MO, selecting A is the preferred option.

**Tableau 7.** *nager* ‘swim’

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. O (E)</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ab. S (A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
Summing up: Auxiliary selection results from the competition of two Mapping Hierarchies, technically Power Hierarchies of mapping constraints themselves derived from harmonic alignment of simple scales referring to lexico-semantic and aspectual features and syntactic configuration. This entails that mapping rules cannot be stated in terms of verb classes, contra Levin & Rappaport Hovav (1995). The constraints proposed here are the same in different languages. What varies is the position of *O constraints relative to the *S Mapping Hierarchy. Thus, variation results from different interactions of the same set of mapping constraints.

4. General Predictions of the OT Analysis

The proposed general OT analysis does not predict total, unconstrained variation in auxiliary selection. Rather, it predicts a very specific typology of languages, sketched out below.

(40) Possible vs. impossible languages:

- Some languages are predicted not to show any split in auxiliary selection. If *O constraints $\gg$ *S constraints, all verb classes select A: e.g. Catalan, Spanish. If all *S constraints $\gg$ *O, all verb classes select E: e.g. Terracinese (a central Italian dialect documented in Tuttle 1986).

- Further languages are predicted to have different cut-offs along the universal hierarchy. Besides Standard Italian and its low cut-off point, Standard French with its high cut-off point, we find Dutch and German with a cut-off point somewhere in between those of Italian and French. In both these Germanic languages, change of location and change of state verb classes select E while the remaining verb classes — continuation of a pre-existing state, existence of state, uncontrolled processes, as well as controlled processes — select A.

  Germanic ranking (preliminary): *S/DIS $\gg$ *S/-HOM $\gg$ *S/+TE $\gg$ *S/+DIR $\gg$ *O $\gg$ *S/+ST $\gg$ *S/-VO $\gg$ *S/MO.

- Languages where existence of state verbs select E but change of state verbs select A are predicted to be impossible. To the best of our
knowledge, this prediction is correct.

- Only 8 languages are predicted possible (= 8 possible positions of *O in the *S hierarchy). If the *S ranking were not universally fixed by virtue of the Power Hierarchy we would expect 8! = 40420 rankings.

(41) Diachronic prediction

Change from a 2-auxiliary system to a 1-auxiliary system should proceed from peripheral to core verb classes and not vice-versa. This prediction is demonstrably correct, based on an independent discussion in Aranovich (2003). In Old Spanish, verbs like trabajar ‘work’ and pecar ‘sin’ never occurred with E. Change from E to A started with the peripheral classes as predicted by our analysis. The first to go were verbs of manner of motion like errar ‘wander’ and verbs of existence of state rastar ‘remain’ (XIV century). Next to change were ‘dynamic verbs of existence and appearance’ (aparecer ‘appear’, desaparecer ‘disappear’, etc.) in the XV century. Morir ‘die’ and ir ‘go’ were the last ones to give up E (XVII century).

5. Indeterminacy in Auxiliary Selection

Recall that Table 2 contains language-particular auxiliary choices annotated with *, indicating that both auxiliaries may in fact be selected. This phenomenon is found with non-core verbs which display a certain amount of fluctuation in their choice of auxiliary. For example, verbs of state in Italian and verbs of appearance in French may select either E or A (see examples in Section 2.3.2).

This type of variation is derivable from the present model, provided it is supplemented by partial constraint ranking, i.e. some indeterminacy in the relative ranking of *O and the *S constraints. By definition, a partial constraint ranking yields a set of rankings (e.g. Anttila 1997, Boersma 1997, Boersma & Hayes 2001, Hagstrom et al. 2001, Legendre et al. 2002, Reynolds 1994). This set of rankings yields potentially different optimal outputs, hence variation in outputs.
Without working out a detailed analysis, we can see how the model works. If *O were to float over four positions towards the top of the *S hierarchy we would obtain a set of four rankings for the class of [-HOM, +TE, +ST] verbs:

\[(42)\] Fixed: 
\[\text{*S} / \text{HOM} \gg \text{*S} / \text{+TE} \gg \text{*S} / \text{+ST}\]
Floating: 
\[\text{*O} \quad \text{*O} \quad \text{*O} \quad \text{*O}\]

\[(43)\] a. \text{*O} \gg \text{*S} / \text{+HOM} \gg \text{*S} / \text{+TE} \gg \text{*S} / \text{+ST}

b. \text{*S} / \text{+HOM} \gg \text{*O} \gg \text{*S} / \text{+TE} \gg \text{*S} / \text{+ST}

c. \text{*S} / \text{+HOM} \gg \text{*S} / \text{+TE} \gg \text{*O} \gg \text{*S} / \text{+ST}

d. \text{*S} / \text{+HOM} \gg \text{*S} / \text{+TE} \gg \text{*S} / \text{+ST} \gg \text{*O}

Different rankings yield different percentages of verbs selecting a particular auxiliary. Specifically, verbs that are -HOM are unaccusative in rankings b, c, d, so 75% of the time. Verbs that are +TE are unaccusative for rankings c, d so 50%; verbs that are +ST are unaccusative only 25% of the time (ranking d). So the more extreme features -HOM and +ST are more homogenously unaccusative or unergative; the middling feature +TE waffles 50/50 between the two. Put another way, there is more indeterminacy in the middle of the range than at the extremes.

6. Conclusion

The Unaccusative Hypothesis has long been a challenge for theories of the lexicon/syntax interface. While there are recurrent patterns both across and within languages the overlap is not complete, as shown by the existence of mismatches for a given test across related languages. This state of affairs is a major challenge to any existing formalization of the lexicon/syntax mapping.

To the extent that the analysis developed here is successful and can be replicated in other languages OT does provide a unique formal resolution of the mapping problem and a successful analysis in turn provides further support for the validity of OT as a theory of syntax and the interfaces.
References


Smolensky, P. 1995. On the internal structure of the constraint component Con of UG, paper presented at University of California, Los Angeles.


