2023. *Proc Ling Soc Amer* 8(1). 5557. https://doi.org/10.3765/plsa.v8i1.5557.

Learning unaccusativity: Evidence for split intransitivity in child Spanish

Victoria Mateu, Laurel Perkins & Nina Hyams*

Abstract. We examine four features of unaccusativity in child-directed and child Spanish to determine what cues children might use to distinguish unaccusative and unergative verbs. Two are cross-linguistic lexico-semantic features: subjects of unaccusatives are patients so we expect more inanimate subjects with unaccusatives; and unaccusatives tend to have an endpoint, hence may occur more frequently with perfective aspect. The other two are language-specific morphosyntactic features: VS order is grammatical with unaccusatives but not unergatives, and many unaccusative verbs allow/require the anticausative *se* clitic. We find all four features robustly in children's input and that even 1-2-year-olds show discriminate use of them.

Keywords. language acquisition; Spanish; split intransitivity; unaccusativity

1. Introduction. To acquire their native language, children must learn the regularities that govern the realization of verb arguments. Transitive verbs take two arguments – a subject and an object – while intransitive verbs take only a subject. Additionally, intransitive verbs split into two subcategories: unergative verbs (e.g., *dance, yell*), whose subject bears an agent-like role, and unaccusative verbs (e.g., *fall, sink*), whose subject exhibits characteristics of an object, such as bearing a theme/patient role (Perlmutter 1978; Burzio 1986; Sorace 2000). Cross-linguistically, this distinction correlates with semantic properties, such as agentivity, but is realized morphosyntactically in variable ways. For instance, in French and Italian each verb type selects a different auxiliary (HAVE vs. BE), and by age 2-3 children use the appropriate auxiliary (Snyder et al. 1995). In this study we examine this phenomenon in Spanish, a language in which this distinction is not marked in the auxiliary system. We ask three questions: (i) are lexico-semantic and morphosyntactic features of split intransitivity in Spanish robustly available in children's input? (ii) are Spanish-acquiring 1-to-5-year-olds sensitive to these features in their own speech? (iii) do they treat "core" and "non-core" unaccusative and unergative verbs alike (Sorace 2000)? This has implications for how split intransitivity is acquired in the face of subtle and variable evidence.

1.1. THE UNACCUSATIVE HYPOTHESIS. Unergative and unaccusative verbs form two distinct categories (e.g., Perlmutter 1978). Semantically, they differ in that the subject of an unergative verbs (e.g., *dance*, *yell*, *cry*) bears the thematic role of the agent, while the subject of unaccusative verbs (e.g., *fall*, *sink*, *die*) acts as a theme/patient, a role typically associated with the object of a transitive clause. Syntactically, unergative verbs select an external argument base-generated in Spec,*v*P as their sole nominal argument, while unaccusative verbs select an internal argument base-generated as the complement of the verb that A-moves to preverbal position (Burzio 1986).

(1) a. The girl sang. (Unergative)
b. The girl fell ~~the girl~~. (Unaccusative)

* We would like to thank our undergraduate Research Assistants at UCLA – April Álvarez, Alexander González, Serena Gutridge, Elaine Heidmann, Shalinee Maitra, and Jose Angelo Urrutia. This study was partly funded by the UCLA Pandemic Relief Fund to VM. Authors: Victoria Mateu, University of California, Los Angeles (vmateu@humnet.ucla.edu), Laurel Perkins, University of California, Los Angeles (perkinsl@humnet.ucla.edu), Nina Hyams, University of California, Los Angeles (hyams@humnet.ucla.edu).

Among the better known syntactic diagnostics of split intransitivity are perfect auxiliary selection (for Italian, French, German, Dutch), the resultative construction (for English), impersonal passivization (for German), and word order (for Italian, Spanish, Mandarin) (Burzio 1986; Huang 1987; Torrego 1989; Levin & Rappaport Hovav 1995; Alexiadou, et al. 2004; a.o.). For example, in languages with two perfect auxiliaries (equivalents of HAVE and BE in English), such as Italian or Dutch, unergative verbs broadly select HAVE, while unaccusatives verbs select BE. In English, the resultative construction is grammatical with transitives and unaccusatives but not unergatives (e.g., *Jake hammered the nail flat*; *The lake froze solid*; **Daniel yelled hoarse*).

However, some studies have argued that a simple unaccusative-unergative split does not reflect the nuances evident in various languages (e.g., Dowty 1991). Sorace (2000) demonstrates that adult speakers of several Western European languages categorically select the auxiliary BE or HAVE with certain classes of verbs while they display variability with others. She thus proposes a gradient hierarchy of unaccusativity based on two lexico-semantic features that characterize split intransitivity – 'agentivity' (active initiator of an event) and 'telicity' (whether the event is bound or has an inherent endpoint). As shown in Table 1, the extreme positions of the hierarchy are occupied with minimally agentive and maximally telic verb classes (core unaccusative) and maximally agentive and minimally telic activities (core unergative).

	Unaccusativity Hierarchy		Inherent agentivity	Inherent telicity
Fall, arrive, come	Change of location	Unaccusative core	Least	Telic
Rise, die, appear	Change of state			
Stay, remain, last	Continuation of state			
Exist, sit, belong	Existence of state			
Shake, cough, shine	Uncontrolled process			
Jump, walk, swim	Motional process	Unergative core		
Work, talk, play	Non-motional process		Most	Atelic

Table 1. Verb categories in Sorace's (2000) Split Intransitivity Hierarchy

Sorace claims that languages differ in the extent to which they base the distinction on one or the other of these two notions, and hence place the boundary between the two verb classes at different points of the hierarchy. The gradient aspect of this approach allows for that variability, not accounted for in binary approaches to the Unaccusative Hypothesis.

To summarize, there is considerable cross-linguistic variation in both the lexico-semantics and the morphosyntax of unaccusativity. Crucially for the learner, not all of these semantic and syntactic features may distinguish the two verb classes with equal precision.

1.2. ACQUISITION OF UNACCUSATIVITY. Research suggests that children establish the properties of the unaccusative/unergative distinction early in various languages (cf. Babyonyshev et al. 2001). Snyder et al. (1995) and Synder and Hyams (2015) examined the spontaneous speech of three French-speaking child and five Italian-speaking children aged 1;05 to 3;05 in the CHILDES database (MacWhinney 2000). All eight children showed near-perfect use of auxiliary selection (HAVE vs. BE) with only six errors being produced out of 326 instances – an unlikely result if the children lacked the adult grammar of unaccusativity.

Other studies show that children are sensitive to the interaction between unaccusativity and word order in their spontaneous speech. Lorusso et al. (2004) found that Italian-speaking children produce more post-verbal subjects with unaccusative verbs but significantly more preverbal

subjects with unergative verbs. Similarly, Friedmann and Costa (2011) tested 1-to-4-year-old children of Hebrew, Palestinian Arabic, European Portuguese, and Spanish using sentence repetition and story retelling tasks. In all four languages, children showed better performance with the VS order than with the SV with unaccusative verbs. A similar task was adopted by Vernice and Guasti (2015), which showed that Italian-speaking children aged 4-5 had better accuracy repeating VS order with unaccusatives than with unergatives.

Comprehension experiments show a similar picture, even in languages like English with subtle cues for split intransitivity (Bunger & Lidz 2004, 2008; Scott & Fisher 2009). For instance, Bunger and Lidz (2004) familiarized English-learning 22- to 25-month-olds to events in which an agent effected a change on a patient (e.g., bouncing a ball), described by novel verbs in potentially unaccusative or transitive frames (e.g., *The ball is pimming* vs. *The girl is pimming the ball*). At test, children saw one video in which the agent performed an action with no change to the object (e.g., a girl patting an immobile ball) and another in which the object underwent a change with no agent (e.g., the ball bouncing on its own). Children looked longer to the change of state video in the unaccusative condition but not the transitive one, showing that they interpreted the inanimate intransitive subject as a patient. Conversely, Bunger and Lidz (2008) found that children interpret animate subjects of intransitives as agents, suggesting that animacy may be used to differentiate unaccusative vs. unergative structures.

In Mandarin, postverbal subjects with the perfective marker *le* are allowed with unaccusative verbs but not unergative verbs. In a visual-fixation task, Wang et al. (2019) presented Mandarin-learning 18- to 20-month-olds with sentences with VS order and *le*. Children looked longer when the sentences contained ungrammatical unergative verbs compared to unaccusative verbs, suggesting that they distinguished between the two intransitive verb types. Similarly, Lin and Deen (2021) conducted a forced-choice task and an acceptability judgment task with Mandarin-speaking children aged 3;0-7;0, manipulating word order and the aspect markers *le* and *zhe* (durative). Overall, children significantly dispreferred VS order with unergative verbs, and the durative aspect with unaccusative verbs. However, the younger children (3-5-year-olds) discriminated between core unaccusative verbs (change of location) and unergative verbs, but not between non-core unaccusative verbs (change of state verbs) and unergative verbs. The authors conclude that the Unaccusativity Hierarchy (Table 1) is in place from birth, and mediates the acquisition of intransitive verbs. In particular, they argue that children are able to identify verbs at the edges of the hierarchy based on their semantic properties (change of location on one end, controlled non-motional process on the other), and thereby bootstrap into the appropriate syntax.

There have been few studies of the acquisition of split intransitivity in Spanish with somewhat conflicting findings. Bel (2003) examined the speech of a Catalan- and a Spanish-speaking child (1;7-2;6) and found they both produced more VS than SV with unaccusatives, but the opposite was true for unergatives. However, Shin (2021) investigated the spontaneous speech of 24 Spanish-speaking children aged 6-8 and she did not find that change of location verbs (i.e., core unaccusative verbs) favored VS order. In the sentence-repetition task by Friedmann and Costa (2011), 12 Spanish-speaking children aged 1;11-3;6 failed to differentiate unaccusative and unergative verbs with respect to word order, showing better performance repeating VS compared to SV sentences for both verb types. They also examined the spontaneous speech of one child aged 1;7-2;7 and he produced more VS than SV orders with both unaccusative and unergative verbs.

In this study, we seek to expand the existing literature on this topic through a larger-scale investigation of the speech of Spanish-speaking children and their caretakers. We consider the pos-

sibility that cues for the categorization of intransitive verbs in Spanish may be less reliable than in previously examined languages and ask if and when children show sensitivity to these cues in their own speech.

1.3. UNACCUSATIVITY IN SPANISH. Unlike Italian or French, Spanish has only one perfect auxiliary (HAVE). However, a handful of diagnostics for split intransitivity have been identified in the literature[1] including the distribution of pre- and postverbal subjects (Torrego 1989; De Miguel Aparicio 1993; Alonso-Cortés 2001; a.o.). In broad-focus contexts, adult Spanish speakers accept both preverbal and postverbal subjects with unaccusative verbs, (2a), but they only accept preverbal subjects with unergative verbs, (2b). Note, however, that post-verbal subjects with unergative verbs are licensed in contrastive focus contexts (3c), among other derived structures (e.g., yes/no-questions, *wh*-questions, relative clauses; Torrego 1984; Zubizarreta 1998; Brucart 2016; Mayoral Hernández 2014; a.o.).

(2) a. {Mamá vino / Vino Mamá}. (Unaccusative)
Mom came came Mom
'Mom came.'

b. {Mamá trabaja / * Trabaja Mamá}. (Unergative)
Mom works works Mom
'Mom works.'

c. Trabaja MAMÁ, no Papá. (Unergative)
works Mom not Dad
'MOM works, not Dad.'

Relatedly, only unaccusative verbs allow for postverbal bare plural subjects, a property shared with objects of transitive verbs (Demonte 1985; Suñer 1982; Torrego 1989; Mendikoetxea 1999; Sanz 2000).

Unaccusative and unergative verbs also pattern differently with regards to the presence of the anticausative *se* clitic. This clitic is sometimes preferred and often required with unaccusative verbs (3a), particularly those that permit a causative-inchoative alternation (Mendikotxea 1999; de Miguel & Fernández-Lagunilla 2000; Cuervo 2014; MacDonald 2017), such as (3b), where the phrase *por sí solo* 'by itself' indicates no agent is implied. Unergative verbs, on the other hand, do not allow anticausative *se* (3c). However, *se* is a highly homophonous clitic in Spanish and appears in a wide variety of other constructions that may involve transitive and even unergative verbs, such as the impersonal (as in 4d), passive, middle, reflexive, reciprocal, dative (when clustered with an accusative clitic), aspectual (telic), and inherent (Klein 1987; de Miguel & Fernández-Lagunilla 2000; MacDonald 2017; a.o.).

(3) a. Juan (se) murió tranquilo. (Unaccusative)
Juan $SE_{ANTICAUS}$ died calm
'Juan died peacefully.'

b. El bosque *(se) quemó por sí solo. (Unaccusative)
the forest $SE_{ANTICAUS}$ burned by self alone
'The forest burned by itself.'

c. El bebé (*se) caminó por primera vez. (Unergative)
the baby $SE_{ANTICAUS}$ walked for first time
'The baby walked for the first time.'

[1] See Roggia (2011) for a complete overview of diagnostics in Spanish.

d. Se camina bien con estos zapatos. (Unergative)
SE$_{IMPERS}$ walks well with these shoes
'One walks well with these shoes.'

Another often cited syntactic test for unaccusativity in Spanish is the participial absolutive construction – only telic unaccusative verbs (i.e., change of location and change of state verbs) are grammatical in this construction, but not unergative or stative unaccusative verbs (i.e., 'continuation of state verbs) (Demonte 1985; De Miguel Aparicio 1993; Campos 1999; Mendikoetxea 1999; Sanz 2000).

As in other languages, there have also been attempts in Spanish to categorize intransitive verbs by their lexical meanings. In terms of theta-role assignment, if the subject bears a role with more proto-patient characteristics, the verb is most likely unaccusative (e.g., *morir* 'die'), and if the subject bears a role with more proto-agent characteristics, the verb is most likely unergative (e.g., *hablar* 'talk'). Verbs that are between the two show variable behavior (Suñer 1982; Mendikoetxea 1999; Alonso-Cortés 2001; Roggia 2011). Accordingly, then, in Spanish, intransitive verbs whose events involve an agent in control of the action (unergative) tend to have preverbal subjects, while those that do not (unaccusative) often have postverbal subjects (López Meirama 1997) (2). The presence of some adverbials, such as *deliberadamente* 'deliberately' or *voluntariamente* 'voluntarily' may also alter the agentivity of the sentence. So, unaccusative verbs such as *entrar* 'to enter', which normally are biased towards VS order, tend to occur with SV order in the presence of an adverbial of this sort (De Miguel Aparicio 1989). Relatedly, some studies of Spanish show that animate subjects of intransitive verbs tend to be preverbal while the inanimate subjects of intransitives are most frequently postverbal, particularly in writing (López Meirama 1997; Rivas 2008). López Meirama (1997) reports that 71% of the animate subjects in his written corpus are preverbal, while only 38% of the inanimate subjects are preverbal. Similarly, Rivas (2008) reports 83% of the animate subjects are preverbal, while only 34% of the inanimate subjects are preverbal.

As discussed by Sorace (2000) (see Table 1), in addition to 'agentivity', the lexical aspect of the verb or its 'telicity' is another lexico-semantic property that characterizes split intransitivity. Telic verbs tend to be unaccusative while atelic verbs tend to be unergative. For example, Spanish *aparecer* 'appear' labels an event with an implicit endpoint, and the verb is unaccusative; conversely, *trabajar* 'work' labels an event without a specified endpoint, and the verb is unergative. In Spanish, grammatical aspect is encoded morphosyntactically by perfective (e.g., preterit *habló* 'spoke', present perfect *ha hablado* 'has spoken') and imperfective morphology (e.g., imperfect *hablaba* 'was speaking', present progressive *está hablando* 'is speaking'). Some studies show that there is a distributional bias in adults and children to use different morphosyntactic aspect markings according to the lexical aspect of the verb – telic verbs tend to occur in the past tense or perfective aspect, while atelic verbs tend to appear in present tense or imperfect aspect (Comrie 1976; Andersen 1986; Shirai 1991; Li & Shirai 2000), as illustrated in (4a-b). Of course, adult Spanish speakers do produce non-prototypical combinations, especially when they include a temporal expression that modifies the boundedness interpretation of the event, as in (4c). However, research shows that non-prototypical combinations are generally dispreferred by adults (Wagner 2009; Domínguez et al. 2013; Quintana Hernandez 2019).

(4) a. La maestra {ha llegado / ?? está llegando}. (Unaccusative)
the teacher has arrived is arriving
'The teacher has arrived / is arriving.'

b. La maestra {??ha hablado / está hablando}. (Unergative)
the teacher has spoken is speaking
'The teacher has spoken / is speaking.'

c. La maestra habl-ó durante dos horas. (Unergative)
the teacher speak-PST.PERF.3SG during two hours
'The teacher spoke for two hours.'

In acquisition studies, Jackson-Maldonado and Maldonado (2001) and Grinstead et al. (2009) have separately shown that Spanish-speaking children perform better in production and comprehension when atelic predicates occur in the simple present (e.g., *Juega*. '(s/he) plays/is playing.') and in the present progressive (e.g., *Está jugando*. 's/he is playing.'), and telic predicates occur in perfective past (e.g., *Tapamos*. 'We closed (it).').

Nevertheless, continuation of state verbs, such as *quedar* 'remain' or *faltar* 'lack' are atelic, but show unaccusative syntax (e.g., preference for VS order, possibility to have postverbal bare plural subject, non-agentive subjects). Therefore, just like the other correlates to unaccusativity in Spanish, the associations among split intransitivity, lexical aspect, and grammatical aspect are not rigid.

1.4. OUR STUDY. In this study, we examine four possible correlates of unaccusativity in child-directed and child-produced Spanish. Two derive from cross-linguistically robust lexico-semantic properties. The first is *subject animacy*: subjects of unergatives tend to name agents and subjects of unaccusatives tend to name patients, so inanimate subjects may be more frequent with unaccusative verbs than unergative verbs. The second is *perfectivity*: the events of unaccusative verbs tend to be telic whereas the events of unergative verbs tend to be atelic, so perfective aspect may be more frequent with unaccusatives than with unergatives. The remaining two correlates are language-specific morphosyntactic features of Spanish. The first language-specific feature is *word order*: adult speakers allow both Subject-Verb and Verb-Subject orders with unaccusatives, but Subject-Verb is preferred with unergatives. The second language-specific feature is the *se* clitic: many unaccusative verbs allow/require anticausative *se*.

However, recall that all of these properties are variably correlated with split intransitivity in Spanish. Although unergative verbs tend to require animate subjects, unaccusative verbs may allow both animate and inanimate subjects. Some (non-core) unaccusative verbs are atelic, and lexical aspect does not always combine with grammatical aspect in a prototypical way. Spanish subjects may be null or appear postverbally in many derived structures (e.g., contrastive focus, yes/no-questions, *wh*-questions, relative clauses). Finally, *se* is a highly homophonous clitic that appears in a number of other unrelated constructions (e.g., reflexive, reciprocal, dative, impersonal, passive, among others). So, these features, particularly the morphosyntactic ones, do not perfectly track verb class.

Given the inconsistent prior acquisition findings in Spanish, it is possible that these signs of unaccusativity in Spanish are not reliable enough to support early acquisition of split intransitivity. However, many past studies had relatively small samples of child and child-directed speech (i.e., N = 1-24 children). Here, we examine a large sample of caregiver and child speech from a large number of children (N = 133 children) in order to obtain a more representative estimate of what children hear and produce. This may help determine whether earlier findings of variable performance in Spanish was due to children's lack of knowledge of split intransitivity, or lack of power to detect that knowledge. We investigate four different potential cues – word order, which has been previously examined (Bel 2003; Friedmann & Costa 2011; Shin 2021), as well as sub-

ject animacy (proxy for agentivity), perfectivity (proxy for telicity), and the anticausative *se* clitic, which have not. Lastly, we examine any potential role of Sorace's (2000) Unaccusativity Hierarchy. The goals of this study are thus to:

(i) provide a large-scale quantitative analysis of these features in child-directed speech and determine if they constitute a reliable indicator of split intransitivity in Spanish.
(ii) provide a large-scale analysis of the speech of young Spanish-speaking children (aged 1-5), to determine if they behave in a way that suggests they are sensitive to each of these cues.
(iii) investigate whether children's distinctive use of these features is first observed for 'core' unaccusative and unergative verbs, and then extends to those categorized as 'non-core' (Sorace 2000), as suggested by Lin and Deen (2021).

2. Corpus study. Data for the study came from three Peninsular Spanish corpora: Aguirre (10,922 child utterances, 23,521 total), BecaCESNo (10,399 child utterances, 22,903 total), and FernAguado (52,344 child utterances, 121,009 total) from the CHILDES database (MacWhinney 2000). Our analysis included spontaneous speech from 133 children aged 1;7-5;11 (median age = 3.52) and their caretakers. Below we describe the extraction and coding procedures that we used to examine the data.

2.1. PROCEDURE. We used the CLAN program (MacWhinney 2000) to select 10 highly frequent unaccusative action verbs and 10 highly frequent unergative action verbs for analysis. We classified these as 'core' vs. 'non-core' using Sorace's (2000) criteria. These are included in (5):

(5) Unaccusative verbs:
a. Core: *caer* 'fall', *entrar* 'enter', *llegar* 'arrive', *mover* 'move'
b. Non-core: *abrir* 'open', *cerrar* 'close', *levantar* 'raise', *morir* 'die', *quedar* 'be left', *romper* 'break'
Unergative verbs:
c. Core: *beber* 'drink', *cantar* 'sing', *comer* 'eat', *funcionar* 'work', *ganar* 'win', *pintar* 'paint'
d. Non-core: *llorar* 'cry', *reír* 'laugh', *saltar* 'jump', *volar* 'fly'

We extracted utterances containing these verbs in both child and adult speech. Many of the verbs listed above allow both transitive and intransitive uses; in order to focus on the characteristics of intransitive clauses, we analyzed only the intransitive uses of these verbs, i.e., when the object was not explicitly produced. We obtained 4,455 adult tokens and 5,229 child tokens in total.

We coded each utterance for the four features under examination: subject animacy, perfectivity, subject position, and *se* clitics. Subjects were coded as animate if they were human, animals, or anthropomorphic toys in cases in which children were engaged in pretend play. Only overt subjects were coded for animacy, as the presence and referents of null subjects could not always be identified reliably. For perfectivity, we analyzed the perfect and preterit as perfective, and present simple, imperfect, progressive forms as imperfective. Subject position was also coded for overt subjects only, and all sentence structures (i.e., declaratives, questions, topicalizations, relative clauses, etc.) were included. For the *se* clitic analysis, we only included utterances in which *se* was used in association with the target verb (as opposed to other verbs in the same utterance) and cases in which *se* did not have a clear argumental role, i.e., we excluded reflexive, reciprocal, and dative uses of *se*. Utterances were hand-coded by a team of trained annotators, with inter-coder agreement above 92% for each feature.

We conducted three main analyses. First, we examined the robustness of each of these features in adult speech vs. child speech, considered as a group. Second, we examined whether children at different ages in our sample showed differences in their sensitivity to these features. For this age comparison, children were divided into 3 terciles: 1-2-year-olds (27 children), 3-year-olds (56 children), and 4-5-year-olds (50 children). Third, we examined whether children and adults' uses of these features differed by each verb's position on the Unaccusative Hierarchy (Sorace 2000); namely, whether the verb was a "core" or "non-core" unaccusative or unergative.

For all comparisons, we only included verbs that had enough usable tokens in all groups to provide the opportunity to observe differences in the use of each feature. That is, for the animacy and word order analyses, we only included verbs for which each group produced at least two tokens with overt subjects, to give us the opportunity to observe differences between animate vs. inanimate, and pre- vs. post-verbal subjects. For the perfectivity analysis, we only included verbs for which each group produced at least two tokens in one of the aspects listed above, to give us the opportunity to observe differences between codable perfective vs. imperfective uses. For any given analysis, these selection criteria excluded no more than four verbs of each type. No verbs were excluded for the *se* analysis.

2.2. RESULTS. In this section we first present the results for our child vs. adult comparisons, followed by child age group, and lastly for verb position in the Unaccusative Hierarchy.

2.2.1. CHILD AND ADULT ANALYSES. Our first analysis examined whether each of the four cues of interest (i.e., subject animacy, perfectivity, word order, and presence of *se* clitic) were distributed differently with unaccusative and unergative verbs in child and adult speech. In Figure 1 we illustrate the distribution of each of the four examined features by verb type and speaker type. For purposes of visualization, for each verb with sufficient codable uses in each speaker group (described above), we calculated the proportion in which each feature was used, and then calculated the average across all verbs. Note we present each of the binary features with the value that favors unaccusative verbs, i.e., inanimate subjects, perfective aspect, VS word order, and presence of *se* clitic. To anticipate our results, we find a significant difference in the distribution of these four features in both adults and children, and in all age groups to the same extent.

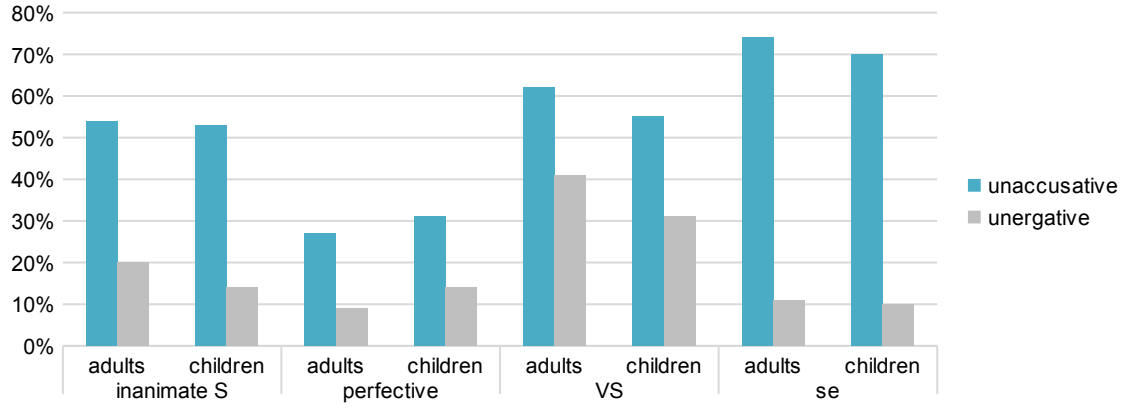

Figure 1. Distribution of features by verb type and speaker type.

We fit a logistic regression model using the *lme4* package (Bates, Mächler, Bolker & Walker 2015) for each of the four features (present, absent), and included verb class (unaccusative, unergative), speaker type (child, adult) and their interaction as predictors. In order to account for the fact that verbs occurred in varying frequencies in the corpus, these analyses used

counts of features for each verb as the dependent measure, rather than the averages illustrated in Figure 1. Model comparisons were conducted using likelihood ratio tests.

In the animacy model, we found a significant main effect of verb class, such that inanimate subjects were more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(35) = 662.3$, $p < .001$). We also found a main effect of speaker type, such that children used slightly more inanimate subjects overall compared to adults ($\chi^2(36) = 1063.13$, $p = .002$).[2] But crucially, we found no interaction between speaker type and verb class ($\chi^2(34) = 659.92$, $p = .122$). Both adults and children used more inanimate subjects with unaccusative verbs than unergative verbs.

In the perfectivity model, we again found a significant main effect of verb class, such that perfective aspect was more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(37) = 918.16$, $p < .001$). We also found a main effect of speaker type, such that children used slightly more perfective aspect overall compared to adults ($\chi^2(38) = 1147.02$, $p < .002$), but we found no interaction between speaker type and verb class ($\chi^2(36) = 917.97$, $p = .666$). Both adults and children used more perfective aspect with unaccusative verbs than unergative verbs.

In the word order model, we again found a significant main effect of verb class, such that postverbal subjects were more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(35) = 112.86$, $p < .001$). We found no main effect of speaker type ($\chi^2(36) = 268.54$, $p = .753$) or interaction ($\chi^2(34) = 111.64$, $p = .27$). Both adults and children produced more postverbal subjects with unaccusative verbs than unergative verbs.

Lastly, in the *se* clitic model, we found a significant main effect of verb class, such that the *se* clitics were more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(37) = 2706.1$, $p < .001$). We also found a main effect of speaker type, such that children used slightly more *se* clitics than adults ($\chi^2(38) = 8736$, $p < .001$), but we again found no interaction ($\chi^2(36) = 2704.8$, $p = .261$). Both adults and children used the *se* clitic more with unaccusative verbs than unergative verbs.

To summarize, verb class (unaccusative or unergative) predicted the distribution of each of our four features of interest: subject animacy, perfectivity, word order, and presence of the *se* clitic. For each feature, both children and adults differentiated between unaccusatives and unergatives in the predicted direction, with no differences between the two speaker groups.

2.2.2. AGE GROUP ANALYSES. Our next set of analyses compared children across ages to determine whether younger children behaved differently from older children. Figure 2 illustrates the average distributions of each of the four examined features by verb type and age group, again including only verbs with sufficient codable uses in each age group. We again fit logistic regression models for counts of each of the four features (present, absent), including verb class (unaccusative, unergative), age group (1-2yo, 3yo, 4-5yo) and their interaction as predictors.

In the animacy model, we found a significant main effect of verb class, such that inanimate subjects were more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(35) = 315.63$, $p < .001$). We also found a main effect of age group, such that children increased their use of animate subjects to more adult-like levels over time compared to the adult results from our first set of analyses (Sec. 2.2.1) ($\chi^2(36) = 498.23$, $p = .001$). We found no interaction between the two ($\chi^2(33) = 314.9$, $p = .69$). All three age groups produced more inanimate subjects with unaccusative verbs than unergative verbs.

[2] This is not illustrated in Figure 1, which represents relative proportions of inanimate subjects by verb class, and not how many inanimate subjects are produced overall by each speaker type; numerically this is higher in children. The same applies to other comparisons.

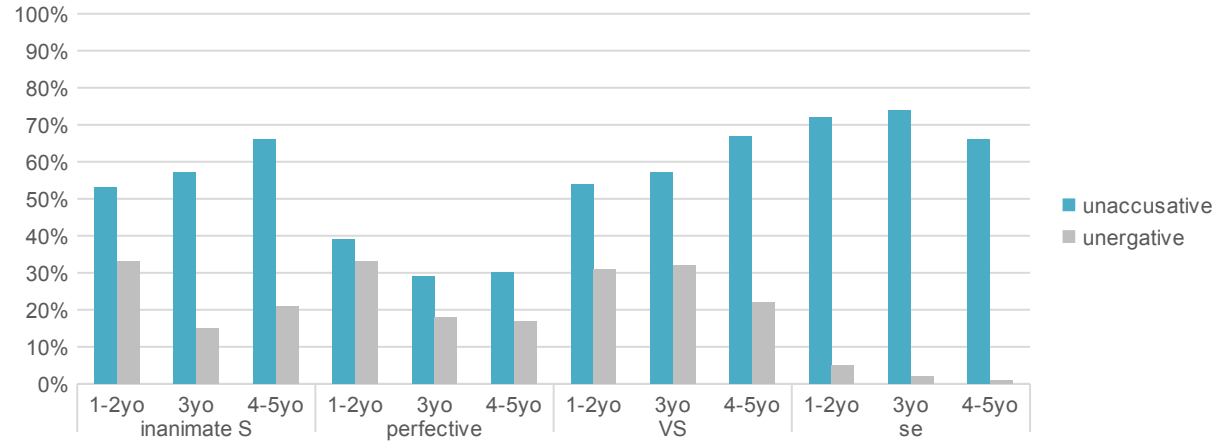

Figure 2. Distribution of features by verb type and age group.

In the perfective model, we found a significant main effect of verb class, such that perfective aspect was more frequently produced with unaccusative verbs than unergative verbs compared to the adult results from our first set of analyses ($\chi^2(50) = 560.91$, $p < .001$). We also found a main effect of age group, such that children increased their use of imperfective aspect to more adult-like levels over time ($\chi^2(51) = 685.48$, $p < .001$) (see Section 2.2.1), and a marginally significant interaction between the two ($\chi^2(48) = 555.37$ $p = .062$). All three age groups produced more perfective marking with unaccusative verbs than unergative verbs, but the difference was slightly greater at younger than older ages.

In the word order model, we found a significant main effect of verb class, such that postverbal subjects were more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(38) = 104.82$, $p < .001$). We found no main effect of age group ($\chi^2(39) = 205.35$, $p = .084$) or interaction between the verb class and age group ($\chi^2(36) = 101.07$, $p = .15$). All three age groups produced more postverbal subjects with unaccusative verbs than unergative verbs.

Lastly, in the *se* clitic model, we found a significant main effect of verb class, such that the *se* clitic was more frequently produced with unaccusative verbs than unergative verbs ($\chi^2(53) = 1357.7$, $p < .001$). We also found a main effect of age group, such that children produced fewer *se* clitics over time ($\chi^2(54) = 4571.6$, $p = .001$), but no interaction between verb class and age group ($\chi^2(51) = 1357.3$, $p = .831$). All three age groups produced more *se* clitics with unaccusative verbs than unergative verbs.

In summary, we found no interaction of verb class and age group in the distribution of the four features of interest. Even the youngest age group differentiated unaccusatives from unergatives in the predicted direction for all four features.

2.2.3. UNACCUSATIVITY HIERARCHY ANALYSES. Our third set of models seeks to address the question of whether these four features distribute differently for core unaccusative and core unergative verbs (Sorace 2000). We fit separate logistic regression models for unaccusative and unergative verb productions by adults and children. The adult models included 'coreness' (core, non-core) as a predictor, and the child models included coreness, age group (1-2yo, 3yo, 4-5yo) and their interaction. Planned comparisons were performed with *z*-ratio tests in *emmeans* when appropriate. A summary of the results is provided in Table 2. To anticipate our findings, we did not find that these four features patterned consistently according to the predictions of the Unaccusativity Hierarchy in child and adult speech.

	1-2yos	3yos	4-5yos	Adults
ANIMACY				
Unaccusative: Core-noncore	-	×	×	×
Unergative: Core-noncore	-	-	-	✓
PERFECTIVITY				
Unaccusative: Core-noncore	-	✓	-	✓
Unergative: Core-noncore	×	×	×	×
WORD ORDER				
Unaccusative: Core-noncore	✓	-	-	-
Unergative: Core-noncore	-	-	-	-
SE CLITIC				
Unaccusative: Core-noncore	-	×	✓	×
Unergative: Core-noncore	-	-	-	✓

Table 2. Results of the Unaccusativity Hierarchy predictions by feature and age group. '×' means a significant effect was found in the unexpected direction, '✓' means a significant effect was found in the expected direction, '-' means no significant effect was found.

In the animacy models, we found an effect of coreness in the expected direction only for unergative verbs in adult speech: adults produced more animate subjects with core vs. non-core unergatives ($\chi^2(8) = 63.014$, $p < .001$). The child model showed no main effects or interaction for unergatives. Both the adult and child models showed an effect in the *unexpected* direction for unaccusatives: more animate subjects used with core vs. non-core unaccusative verbs (adults: $\chi^2(8) = 229.8$, $p < .001$; children: $\chi^2(17) = 171.36$, $p < .001$). We additionally found an interaction with age ($\chi^2(15) = 163.87$, $p = .023$), such that this unexpected core/non-core difference for unaccusatives was significant in the 3- to 5-year-olds only (z-ratios > 5.86, ps < .001).

In the perfectivity models, we found an effect of coreness in the expected direction only for unaccusative verbs: both children and adults used perfective aspect markers more with core unaccusative verbs (adults: $\chi^2(8) = 245.86$, $p < .001$; children: $\chi^2(26) = 347.91$, $p < .001$). This effect interacted with age ($\chi^2(24) = 327.47$, $p < .001$), such that 3-year-olds showed the expected difference between core/non-core unaccusatives ($p < .001$), but the other age groups did not. For unergative verbs, we found a main effect of coreness in the *unexpected* direction— more use of perfective aspect with core unergative verbs (adults: $\chi^2(8) = 98.047$, $p < .001$; children: $\chi^2(20) = 150.3$, $p = .003$). No interactions by age were found with unergatives.

In the word order models, we found an interaction of coreness and age for unaccusatives in child speech ($\chi^2(15) = 68.716$, $p = .001$). Planned comparisons revealed a difference in the expected direction only for the 1-2-year-olds (z-ratio = 3.18, $p = .002$): more use of VS word order with core unaccusatives. We found no significant effects in the adult models, and no significant effects or interactions for unergative verbs, for either children or adults.

In *se* clitic models, we found a main effect of coreness in the expected direction only for unergative verbs in adult speech: fewer uses of *se* with core unergative verbs ($\chi^2(8) = 68.954$, $p < .001$). The child model showed no main effects or interactions for unergatives. For unaccusative verbs, we found an effect of coreness in adult speech in the *unexpected* direction: they used *se* more often with non-core unaccusatives ($\chi^2(8) = 1222.8$, $p = .003$). The child model for unaccusatives revealed an interaction with coreness ($\chi^2(24) = 1309.9$, $p < .001$), such that 3-year-

olds, like adults, produced more *se* with non-core unaccusatives (z-ratio = -2.55, p = 0.01), while 4-5-year-olds produced more *se* with core unaccusatives (z-ratio = 3.01, p = 0.002).

To summarize, we did not find evidence that the Unaccusativity Hierarchy plays a consistent role in the distribution of animate subjects, perfective aspect, word order, or *se* clitics in child or adult Spanish. We did not always find a difference between core and non-core unaccusative and unergative verbs. When we did find a significant difference, it did not appear to follow a logical developmental trend, or it was counter to the predictions of the Unaccusativity Hierarchy.

3. Discussion. Earlier findings showed inconsistent evidence for knowledge of split intransitivity in Spanish learners (Bel 2003; Friedmann & Costa 2011; Shin 2021). Our study sought to illuminate this question through a much larger-scale investigation of child and adult speech. We examined four cues to unaccusativity in child-directed and child-produced Spanish. Two are lexico-semantic properties that are cross-linguistically common: agentivity (with animacy as proxy) and telicity (with perfective aspect as proxy). Two are language-specific morphosyntactic features: postverbal subjects and presence of anticausative *se*. Subjects in Spanish may be null or appear postverbally in many derived constructions, such as focus constructions, questions, and relative clauses, and the clitic *se* is often found in a number of other constructions that are not unaccusative. Therefore, these cues do not perfectly track unaccusativity.

Results from our extensive corpus analysis (9,684 analyzed utterances) show that unaccusative verbs have a significantly higher rate of inanimate subjects, perfective aspect marking, postverbal subjects, and nonargumental *se* clitics in both adult and child spontaneous speech. Additionally, we find that even the youngest children examined in the study (aged 1;7-2;11) are highly sensitive to the cues that differentiate the unaccusative/unergative classes in Spanish. Our results demonstrate that children are attuned to both cross-linguistically common lexico-semantic features and highly variable language-specific morphosyntactic cues very early in development.

Interestingly, our results did not show that verbs at the edges of the Unaccusative Hierarchy (Sorace 2000) behaved in a more canonical way than those that lie in between: overall, core unaccusative verbs did not exhibit higher rates of inanimate subjects, perfective aspect marking, postverbal subjects and *se* clitics compared to non-core unaccusatives, and vice-versa for unergative verbs. This is in contrast to experimental findings in Lin and Deen (2021), in which Mandarin-speaking children first discriminated only core unaccusatives from unergatives, and only later discriminated both core and non-core unaccusative verbs from unergative verbs. It is unclear whether this is because the Unaccusative Hierarchy fails to play as clear a role in (Spanish) language acquisition, or whether the "core" vs. "non-core" verbs in our sample were in some way idiosyncratic. We leave this question for future studies.

Overall, our results suggest that Spanish-speaking children have acquired the appropriate unaccusative/unergative categorization by age 1;7 and use the distinctive features of Spanish to differentiate these verb types in an adult-like fashion. However, this leaves open the question of whether children's early sensitivities reflect knowledge of the abstract grammatical properties of these two verb classes, or whether it is a product of learning the distributional properties of verbs in a lexically-specific manner (Tomasello 1992). These findings therefore invite further experimental work to test whether children at early stages of verb learning use these cues to unaccusativity to inform their learning of new intransitive verbs.

References

Alexiadou, Artemis, Elena Anagnostopoulou & Martin Everaert (eds.). 2004. *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199257652.001.0001.

Alonso-Cortés, Ángel. 2001. Rasgos activos y ergativos del español. *Verba* 28. 187–206.

Andersen, Roger W. 1986. The need for native language comparison data in interpreting second language data. Forum Lecture. TESOL Summer Institute, University of Hawaii at Manoa.

Babyonyshev, Maria, Jennifer Ganger, David Pesetsky & Kenneth Wexler. 2001. The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32(1). 1–44. https://doi.org/10.1162/002438901554577.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). 1–48. https://doi.org/gcrnkw.

Bel, Aurora. 2003. The syntax of subjects in the acquisition of Spanish and Catalan. *Probus* 15. 1–26. https://doi.org/10.1515/prbs.2003.003.

Brucart, José María. 2016. Oraciones de relativo. In Javier Gutiérrez-Rexach (ed.), *Enciclopedia de lingüística hispánica*, 722–736. London: Routledge. https://doi.org/d766.

Bunger, Ann & Jeffrey Lidz. 2004. Syntactic bootstrapping and the internal structure of causative events. In Alejna Brugos, Linnea Micciulla & Christine E. Smith (eds.), *Proceedings of the 28th annual Boston University Conference on Language Development*, 74–85. Somerville, MA: Cascadilla Press.

Bunger, Ann & Jeffrey Lidz. 2008. Thematic relations as a cue to verb class: 2-year-olds distinguish unaccusatives from unergatives. *University of Pennsylvania Working Papers in Linguistics* 14(1). 4.

Burzio, Luigi. 1986. *Italian syntax: A government-binding approach*. Dordrecht: Springer. https://doi.org/10.1007/978-94-009-4522-7.

Campos, Héctor. 1999. Transitividad e intransitividad. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua Española*, 1519–1574. Madrid: Espasa Calpe.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Cuervo, María Cristina. 2014. Alternating unaccusatives and the distribution of roots. *Lingua* 141, 48–70. https://doi.org/10.1016/j.lingua.2013.12.001.

De Miguel Aparicio, Elena. 1993. Construcciones ergativas e inversión en la lengua y la interlengua española. In Juana Muñoz Liceras (ed.), *La Lingüística y el análisis de los sistemas no nativos.* 178–195. Ottawa: Dovehouse Editions.

De Miguel, Elena & Marina Fernández Lagunilla. 2000. El operador aspectual *se*. *Revista Española de Lingüística* 30(1). 13–44.

Demonte, Violeta. 1985. Papeles temáticos y sujetos sintácticos en el sintagma nominal. *Rivista di Grammatica Generativa* 9-10. 265–331.

Domínguez, Laura, Nicole Tracy-Ventura, María J. Arche, Rosamond Mitchell & Florence Myles. 2013. The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology *Bilingualism: Language and Cognition* 16(3). 558–577. https://doi.org/f4x8kj.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. https://doi.org/10.2307/415037.

Friedmann, Naama & João Costa. 2011. Acquisition of SV and VS order in Hebrew, European Portuguese, Palestinian Arabic, and Spanish. *Language Acquisition* 18(1). 1–38. https:// doi.org/10.1080/10489223.2011.530507

Grinstead, John, Teresa Pratt & Dan McCurley. 2009. Comprehension of prototypical tense and aspect combinations in child Spanish. *Studies in Hispanic and Lusophone Linguistics* 2(2). 435–450. https://doi.org/10.1515/shll-2009-1059.
Huang, C.-T. James. 1987. Existential sentences in Chinese and (in)definiteness. In Eric Reuland & Alice ter Meulen (eds.) *The representation of (in)definiteness*, 226–253. Cambridge, MA: MIT Press.
Jackson-Maldonado, Donna & Ricardo Maldonado. 2001. Determinaciones semánticas de la flexión verbal en la adquisición temprana del español. In Cecilia Rojas-Nieto & Lourdes De León Pasquel (eds.), *La adquisición de la lengua materna: Español, lenguas mayas, euskera*. 165–200. Mexico City: UNAM-CIESAS.
Klein, Philip W. 1987. Syntax and semantics of Spanish *se*. *Hispanic Journal* 9(1). 149–162.
Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-lexical semantics interface.* Cambridge, MA: MIT Press.
Li, Ping & Yasuhiro Shirai. 2000. *The acquisition of lexical and grammatical aspect* (Studies in Second Language Acquisition 16). De Gruyter Mouton. https://doi.org/dk6c2q.
Lin, Kai-ying & Kamil Ud Deen. 2021. Unaccusativity in Mandarin child language. In Danielle Dionne & Lee-Ann Vidal Covas (eds.), *Proceedings of the 45th Annual Boston University Conference on Language Development*, 486–498. Somerville, MA: Cascadilla Press.
López Meirama, Belén. 1997. *La posición del sujeto en la cláusula monoactancial en español*. Santiago de Compostela: Universidade de Santiago de Compostela.
Lorusso, Paolo, Claudia Caprin & Maria Teresa Guasti. 2005. Overt subject distribution in early Italian children. In Alejna Brugos, Manuella R. Clark-Cotton & Seungwan Ha (eds.), *A Supplement to the Proceedings of the 29th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
MacDonald, Jonathan E. 2017. Spanish aspectual *se* as an indirect object reflexive: The import of atelicity, bare nouns, and leísta PCC repairs. *Probus* 29(1). 73–117. https://doi.org/j62v.
MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk, Volume 1: Transcription format and programs*. New York: Psychology Press.
Mayoral Hernández, Roberto. 2014. Subject position in Spanish: a study of factor interactions with prototypical verbs. In Andrés Enrique-Arias, Manuel Gutiérrez, Alazne Landa & Francisco Ocampo (eds.), *Perspectives in the study of Spanish language variation: Papers in honor of Carmen Silva-Corvalán,* 111–136. Santiago de Compostela: Universidade de Santiago de Compostela.
Mendikoetxea, Amaya. 1999. Construcciones inacusativas y pasivas. In Ignacio Bosque & Violeta Demonte (eds.) *Gramática descriptiva de la lengua española*, 1575–1629. Madrid: Espasa Calpe.
Perlmutter, David M. 1978. Impersonal passives and the Unaccusative Hypothesis. *Berkeley Linguistics Society (BLS)* 4. 157–190. https://doi.org/10.3765/bls.v4i0.2198.
Quintana Hernández, Lucía. 2019. The effects of telicity, dynamicity and punctuality in L2 acquisition of Spanish preterit and imperfect. *Dutch Journal of Applied Linguistics* 8(1). 100–116.
Rivas, Javier. 2008. La posición del sujeto en las construcciones monoactanciales del español: Una aproximación funcional. *Hispania* 91(4), 897–912.
Roggia, Aaron B. 2011. *Unaccusativity and word order in Mexican Spanish: An examination of syntactic interfaces and the split intransitivity hierarchy*. State College, PA: The Pennsylvania State University dissertation.

Sanz, Montserrat. 2000. *Events and predication: A new approach to syntactic processing in English and Spanish*. Amsterdam: John Benjamins.
Scott, Rose M. & Cynthia Fisher. 2009. Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes* 24(6). 777–803. https://doi.org/10.1080/01690960802573236.
Shin, Naomi L. 2021. Acquiring constraints on variable morphosyntax: Subject–verb ~ verb–subject order in child Spanish. In Manuel Díaz-Campos (ed.) *The Routledge handbook of variationist approaches to Spanish*, 425–436. London: Routledge. https://doi.org/j62w.
Shirai, Yasuhiro. 1991. *Primacy of aspect in language acquisition: Simplified input and prototype*. Los Angeles, CA: UCLA dissertation.
Snyder, William & Nina Hyams. 2015. Minimality effects in children's passives. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond. Studies in honour of Adriana Belletti*, 343–368. Philadelphia: John Benjamins. https://doi.org/10.1075/la.223.
Snyder, William, Nina Hyams & Paola Crisma. 1995. Romance auxiliary selection with reflexive clitics: Evidence for early knowledge of unaccusativity. In Eve Clark (ed.) *Proceedings of the 26th Annual Child Language Research Forum*, 127–136. Stanford, CA: CSLI.
Sorace, Antonella. 2000. Gradients in auxiliary selection with intransitive verbs. *Language* 76(4). 859–890. https://doi.org/10.2307/417202.
Suñer, Margarita. 1982. *Syntax and semantics of Spanish presentational sentence-types*. Washington DC: Georgetown University Press.
Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1). 103–129.
Torrego, Esther. 1989. Unergative-unaccusative alternations in Spanish. *MIT Working Papers in Linguistics* 10, 253–272.
Vernice, Mirta & Maria Teresa Guasti. 2015. The acquisition of SV order in unaccusatives: Manipulating the definiteness of the NP argument. *Journal of Child Language* 42(1). 210–237. https://doi.org/10.1017/S0305000913000536.
Wagner, Laura. 2009. I'll never grow up: Continuity in aspect representations. *Linguistics* 47(5). 1051–1074. https://doi.org/10.1515/LING.2009.037.
Wang, Ziqi, Xiaolu Yang & Rushen Shi. 2019. Mandarin-speaking toddlers' acquisition of unaccusativity. In Megan M. Brown & Brady Dailey (eds.), *Proceedings of the 43rd Boston University Conference on Language Development*, 705–715. Somerville, MA: Cascadilla Press.
Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge, MA: MIT Press.