

# (Im)possible Idioms in a Minimalist Theory of Argument Structure: Insights from Spanish

Enrique Merino Hernández

## 1. Introduction

The study of idioms has been of interest to linguists coming from a wide variety of theoretical backgrounds. In more syntax-oriented approaches to language, like generative grammar, a great effort has been made to unravel the structural conditions under which idiomatic readings can arise. Bruening (2010), for example, puts forth the linguistic principles in (1) to account for the range of possible idioms.

- (1) a. *The Principle of Idiomatic Interpretation* (Bruening 2010: 532, (24))  
X and Y may be interpreted idiomatically only if X selects Y.
- b. *Constraint on Idiomatic Interpretation* (Bruening 2010: 532, (25))  
If X selects a lexical category Y and X and Y are interpreted idiomatically, all of the selected arguments of Y must be interpreted as part of the idiom that includes X and Y.
- c. Lexical categories are V, N, A, Adv. (Bruening 2010: 532, (26))

Although it must be acknowledged that Bruening's proposed constraint correctly captures the relevant empirical phenomena, i.e., it is descriptively accurate (but see Bruening 2020 for qualification), it lacks explanatory power in the sense that the question as to why such constraint should be in order at all is left without a response. The constraint also relies on the somewhat blurry distinction between functional and lexical categories; this distinction, at least in its traditional guise, has been contested by antilexicalist approaches to grammar like Distributed Morphology (Halle & Marantz 1993 and subsequent work; see especially Marantz 1997), where the notion of category is not a primitive of the theory.<sup>1</sup> Furthermore, given recent developments within Chomskyan linguistics, Bruening's proposal and the like do not seem plausible from an evolutionary perspective, as they complicate the internal structure of UG by adding specifications about what can be an idiom and what cannot; this poses a formidable challenge for biological theories of language.

Alternatively, the present short paper argues that idiom size is constrained by language-design issues pertaining to the so-called 3<sup>rd</sup> factor (Chomsky 2005), while keeping UG as bare as empirically possible (Chomsky 2007). More specifically, evidence is provided that only those elements that belong to the lower VP phase can be a fixed part in an idiom (Svenonius 2005); this excludes agents (Marantz 1997, Harley & Stone 2013), and, as will be shown, high applied arguments too.

## 2. External vs. internal argument asymmetries in idioms: The case of agents

Marantz (1984) made the influential observation that external arguments, as opposed to internal ones, cannot condition the interpretation of the verb by which they are selected for. In (2), the truth conditions

---

\* Center for Theoretical Linguistics (Autonomous University of Barcelona), enrique.merino@uab.cat. The present investigation has been possible thanks to a predoctoral research grant from the Basque Government and the PID2021-123617NB-C41 project (*Microparameters and Networks in Romance Variation*, MINERVA) granted by the Spanish Ministry of Science and Innovation. Special thanks go to Jaume Mateu and Isabel Oltra-Massuet.

<sup>1</sup> In Distributed Morphology the atoms of syntactic computation are divided into two types: abstract morphemes and roots. By definition, roots lack a category. Roots are categorized in the syntax by merging with dedicated functional heads (*v*, *n*, *a*). Consequently, *stricto sensu* there is no such thing as a divide between lexical and functional categories, but rather between functional elements (i.e., abstract morphemes) and roots, which are categoryless.

of the verb *kill* depend on the denotation of the internal argument; in (3), however, the external argument does not have an effect on the truth conditions of the verb. Based on this empirical generalization, Kratzer (1996) proposed her Voice hypothesis: external arguments are introduced by an independent functional head in the syntax, Voice, whereas internal arguments are directly selected for by the lexical core of the verb (4). Semantically, the external argument does not combine with the verb by means of Functional Application, but rather via a new operation which Kratzer dubbed Event Identification. The explanation to the internal *vs.* external argument asymmetry discovered by Marantz receives a semantic implementation in Kratzer's work: only those elements that directly compose with the verb via Functional Application, that is, internal arguments, can activate special, context-dependent interpretations on it.

- (2) Kill a cockroach / a conversation / an evening watching TV.  
 (3) Harry / Everyone / Silence killed NP.  
 (4) [<sub>VoiceP</sub> External Argument [<sub>Voice'</sub> Voice [<sub>VP</sub> V [ Internal Argument ]]]]

Marantz (1997), however, offers an alternative proposal that does not rely on semantic composition. Assuming a bipartite division of the VP, he hypothesizes that the head introducing agents delimits a locality domain for special interpretations. An implication of this hypothesis is that there should be no idioms of the sort of *The dog bites DP*, where the agent is fixed and the internal argument is allowed to vary freely. In fact, it has been argued that such idioms are non-existent cross-linguistically (see Harley & Stone 2013 for discussion of some apparent counterexamples). Under such an account, then, idiomaticity is constrained by principles of locality, and not by strictly semantic factors. Given current advancements within generative grammar, it seems natural to couple Marantz's proposed locality domain with the minimalist notion of a phase (Chomsky 2000, 2001 and subsequent work); this idea is especially appealing because the head introducing agents has been argued to be a phase head on independent grounds. All in all, if idioms cannot span two phases, as Svenonius (2005) has explicitly proposed, and if the head introducing agents (Voice, in our terms) is a phase head, then the 'No Agent Idioms' generalization follows naturally. With this background in mind, the present paper explores the behavior of different argument introducers in Spanish idiomatic expressions involving dative DPs to conclude that Voice is not the only head that sets the limit to special meanings; rather, a more general constraint is in order that gives even more credit to the phase-based account of idiom size.

### 3. Theoretical background

#### 3.1. Distributed Morphology

Distributed Morphology (Halle & Marantz 1993 and subsequent work) is a theory of grammar according to which the only device of human language with generative capacity is the syntax. Thus, Distributed Morphology denies the existence of a generative lexicon altogether and assumes that morphemes, in an abstract sense, are directly combined in the syntax by means of a binary operation referred to as Merge (Chomsky 1995). Consequently, all linguistic structure is phrasal in essence. In such a framework, the study of idioms becomes of great interest, because as soon as the existence of a generative lexicon is rejected, the idea has to be embraced that idioms are formed directly in the syntax. Evidence has been provided from different empirical domains that that is in fact the case, either because (i) idioms are syntactically well-behaved objects, or (ii) idiomaticity is subject to syntactic constraints. Precisely, the main goal of this paper is to provide further evidence in favor of (ii).

#### 3.2. The pieces of argument structure

In what follows, I will assume, together with much work within Distributed Morphology, that there exist three classes of building blocks of argument/event structure. Roots function as a repository of real-world/encyclopedic information; it is a widely accepted fact that roots do not have a category (see fn. 1).<sup>2</sup> Little *v* heads serve to categorize roots in the syntax and to introduce an event variable at LF. Finally, argument introducers, such as Voice or Applicative, serve to add extra participants to the event.

<sup>2</sup> The question of whether roots contain grammatical information or not is highly contested. See Harley (2014a, b) and the references there for a general overview.

### 3.3. Datives in Spanish

I adopt Cuervo's (2003, 2020) proposal that all dative DPs in Spanish are introduced in the specifier position of an applicative head. Based on their attachment site, Pyllkkänen (2002/2008) classifies applicative heads into two types: Low Applicative and High Applicative. This distinction has semantic implications. Low Applicative, which is embedded within *v*, syntactically relates two DPs (5). Semantically, it denotes a relation between two entities. The dative DP in a Low Applicative construction can be variously interpreted as a recipient (6a), a source (6b), or an unaffected possessor (6c).

- (5) [<sub>vP</sub> v [<sub>LowAppIP</sub> DP [<sub>LowAppI</sub> LowAppl [DP]]]]
- (6) a. El profesor le dio las llaves al alumno.  
 the teacher CL.DAT gave the keys [the student].DAT  
 'The teacher gave the keys to the student.'
- b. El profesor le quitó las llaves al alumno.  
 the teacher CL.DAT took the keys [the student].DAT  
 'The teacher took the keys from the student.'
- c. El detective le vio la cara al sospechoso.  
 the detective CL.DAT saw the face [the suspect].DAT  
 'The detective saw the suspect's face.'

On the other hand, High Applicative takes a *vP* as its complement and introduces a DP in its specifier position (7). Semantically, it relates the entity denoted by the DP to the event(uality) denoted by the *vP*. The interpretation that the high applied argument receives is pretty underspecified, and it can range from affectee (8 and 9) to accidental causer (9); the key point is that in all the cases the argument in question is related to the whole event, and not to some other participant directly (note that in the examples below, for example, the dative arguments are not necessarily the possessors of the theme).

- (7) [<sub>HighAppIP</sub> DP [<sub>HighAppI</sub> HighAppl [<sub>vP</sub> v [VP]]]]
- (8) Juan le ha cerrado la puerta a María.  
 Juan CL.DAT has closed the door María.DAT  
 'Juan has closed the door on María.'
- (9) A Juan se le ha cerrado la puerta.  
 Juan.DAT SE CL.DAT has closed the door  
 'The door has closed on Juan. / Juan has accidentally closed the door.'

With this background in mind, let us now examine the behavior of dative DPs in Spanish idioms.

## 4. Dative DPs in Spanish idioms

### 4.1. The data

Spanish is rich in idiomatic expressions involving a dative DP. However, there is a difference between those idioms in which the dative is fixed, i.e., with the encyclopedic information associated with the dative listed as part of the idiom, as in the idioms in (10); or open, i.e., being the free variable in the idiom (represented by an X), as in the idioms in (11).

- (10) a. Vender-(le) el alma al diablo.  
 sell-CL.DAT the soul [the devil].DAT  
 'To sell one's soul to the devil.'
- b. Pedir-(le) peras al olmo.  
 ask-CL.DAT pears [the elm].DAT  
 'To expect something impossible to happen (lit. ask pears from the elm).'
- c. Ver-(le) las orejas al lobo.  
 see-CL.DAT the ears [the wolf].DAT  
 'To start feeling scared about a situation (lit. see the wolf's ears).'

- (11) a. *Ír-se-le a X el santo al cielo*  
 go-SE-CL.DAT X.DAT the saint to-the sky  
 ‘To become distracted (lit. the saint goes to heaven on X).’
- b. *Encendér-se-le a X la bombilla.*  
 light.up-SE-CL.DAT X.DAT the lightbulb  
 ‘To have a brilliant idea (lit. the lightbulb lights up for X).’
- c. *Cerrar-se-le a X todas las puertas.*  
 close-SE-CL.DAT X.DAT all the doors  
 ‘To be left out with no resources (lit. all doors closed on X).’

A closer examination of the sets of idioms in (10) vs. those in (11) brings to light a series of interesting differences:

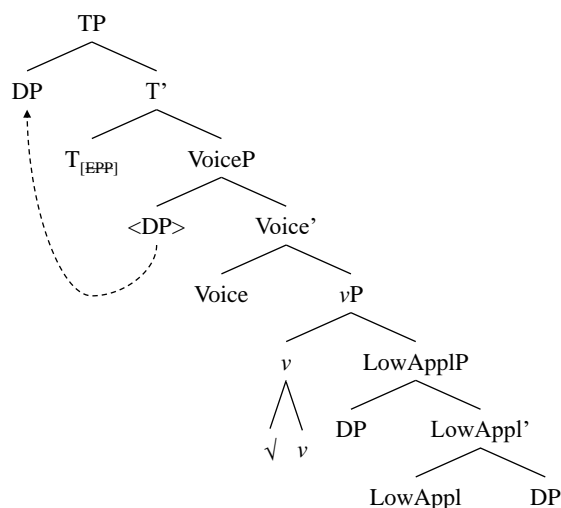
- i. All the idioms in (11) involve *se*-marked unaccusatives, whereas the idioms in (10) include (di)transitive verbs.
- ii. In the idioms in (11) the nominative subject is a fixed part of the idiom (e.g., *el santo* in (11a)); these are in actual fact internal arguments based-generated in object position (Perlmutter 1978, Burzio 1986). On the other hand, the idioms in (10) do not involve fixed nominative subjects, arguably because these are true external arguments (see Section 2 above).
- iii. In the idioms in (10), the dative clitic *le* is optional; this is not the case for the idioms in (11).

#### 4.2. Discussion

In what follows, I will assume, together with much work in Distributed Morphology, that the morphosyntactic properties of idioms are the same as their literal counterpart’s (Marantz 1997; McGinnis 2002, 2005; Mateu 2023; among others). With this in mind, and building on Pykkänen’s (2002/2008) and Cuervo’s (2003) proposals about applicatives, the following conclusions can be drawn.

On the one hand, the idioms in (10) involve a Low Applicative. Evidence in favor of this comes from the fact that in these cases the dative DP denotes a recipient (10a), a source (10b), or an unaffected possessor (10c), i.e., in general, the dative DP stands in a relation with the theme. Moreover, the dative clitic can be dropped; if Pineda (2013) is on the right track, in Low Applicative constructions spell-out of the dative clitic is optional.

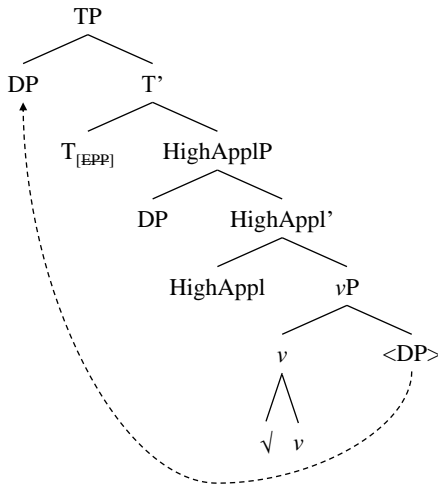
(12) Syntactic derivation for the transitive idioms in (10)



On the other hand, the idioms in (11) involve a High Applicative. These idioms include *se*-marked unaccusatives, which, as argued for by Cuervo (2003), are incompatible with a Low Applicative. Semantically, the dative DP is related to the whole event: the dative denotes the entity which is affected by the change-of-state/location event. Furthermore, there is no optionality regarding the presence of the

dative clitic: it is obligatory. Adopting Pineda's proposal, this can be accounted for if these idioms do not involve a Low Applicative.

(13) *Syntactic derivation for the unaccusative idioms in (11)*



Given these facts, it can be concluded that there is a structural asymmetry between the datives in (10) and the datives in (11).

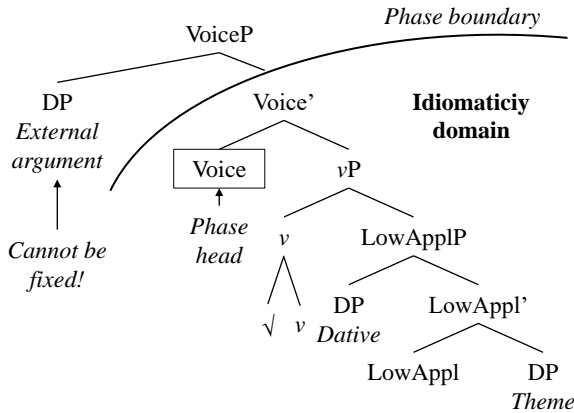
## 5. Accounting for the facts

If the structural asymmetry proposed in the previous section is on the right track, the fact that the datives in the idioms in (11) are never fixed can receive a principled explanation in terms of phases. The guiding hypothesis here, following Svenonius (2005), is that an idiom cannot contain listed information belonging to two different phases. Intuitively, as Harley & Stone (2013: 254-255) put it, “at the point of interpretation, lexical entries must be accessed to construct the semantic representation, and only information within the phase that is being interpreted could condition the selection of the appropriate truth-conditions needed for this representation.”

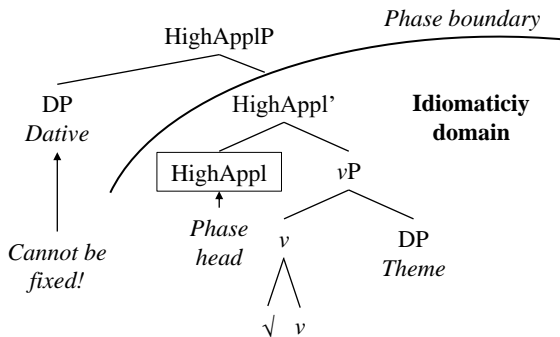
McGinnis (2001) argues that there is an asymmetry in the phasal status of applicative heads. More specifically, she shows that applicatives that relate an entity to an event (High Applicative, in Pylkkänen's terminology) are phase heads. On the other hand, applicatives that relate two entities (i.e., Low Applicative) do not head a phase. Den Dikken (2023) reaches a similar conclusion when studying the behavior of case assignment in unaccusatives. The dative asymmetry in the idioms discussed in this paper can then be elegantly accounted for by resorting to the notion of phase. Because Low Applicative does not define a phase, the fact that both the dative and the theme in the idioms in (10) is fixed is not surprising: the limit to idiomatity will be demarcated by the higher phase head, i.e., Voice. This means that in these idioms the nominative subject, being a true external argument, cannot be fixed, as the ‘No Agent Idioms’ generalization predicts. The prediction is borne out. On the other hand, because High Applicative does define a phase, it sets the limit to idiomatity. This means that the dative in the idioms in (11) cannot be fixed, as it and the fixed lexical material below it belong to two different phases. The phase-based hypothesis of idiom formation therefore makes accurate predictions in this domain.

If all the fixed idiomatic material has to belong to the same phase, then the Spanish data show that the High Applicative head must be interpreted together with its complement, because the presence of a dative is required so that the idiomatic reading can arise. In other words, the presence of a dative element is obligatory, but the specific content of the dative is allowed to vary; this indicates that it is the head selecting the dative that is a fixed part of the idiom, and not the dative DP *per se*. From this it follows that the phase head and its complement have to be interpreted together at LF; Transfer thus applies to HighAppI'. The tree diagrams in (14) and (15) capture the proposals made in this section.

## (14) Proposal for the idioms in (10)



## (15) Proposal for the idioms in (11)



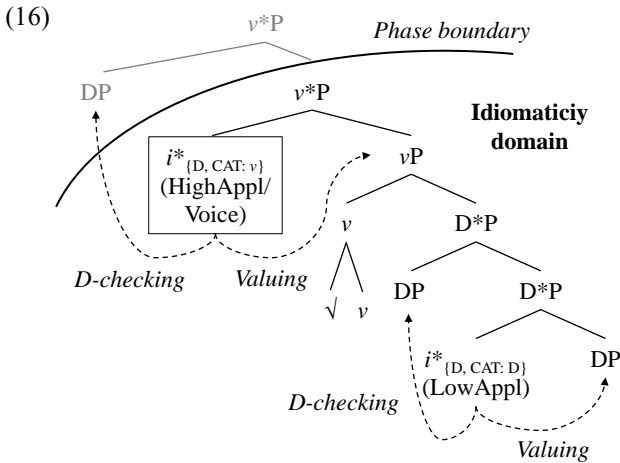
## 6. Formalizing the proposal

At this point, the question arises of whether the ‘No Agent Idioms’ generalization and the present discovery about datives can be formally unified. I propose that the two constraints can be given a common explanation by making use of Wood & Marantz’s (2017) [W&M] minimalist theory of argument structure.

W&M seek to reduce the inventory of functional items by subsuming categories like Voice, Appl, and *p* under the same argument-introducing head  $i^*$ , which has the following syntactic properties:

- It comes with an unvalued category feature and an uninterpretable D feature.
- The unvalued category feature receives a value from the category of the element in  $i^*$ -complement position.
- The D feature gets checked by merging something of category D in the specifier position of  $i^*$ ; this formal implementation accounts for the argument-introducing capacity of  $i^*$ .
- $i^*$  closes off the extended projection of the first element it merges with.

If we assume that argument introducers are all the same head, then the distinction between Voice and High Applicative is not lexically specified, and consequently the theory does not have the means to resort to a putative lexical distinction to define the phasehood of these heads. However, we can characterize a phase head at the VP level simply as  $i^*$  valued as  $v$ ; this is essentially the argument introducer that closes off the extended projection of the verb. Because both Voice and High Applicative are argument introducers merged right on top of a vP, the formalization correctly predicts that these two are phase heads. On the other hand, because Low Applicative is an argument introducer merged below  $v$ , its categorial status cannot possibly be  $v$ . The present proposal, then, correctly captures the fact that Low Applicative is not a phase head. The proposal is schematized in the tree diagram below.



All in all, W&M's minimalist theory of argument structure provides us with the tools to unify the phasal status of heads like Voice and High Applicative, which is a welcome result given the present discovery about the unavailability of high applied datives in Spanish idioms.

## 7. Conclusions

This paper has discussed a dative asymmetry in Spanish idioms that, to the best of my knowledge, had remained unnoticed in the literature. More specifically, it has been shown that high applied arguments, together with agents, cannot be a fixed part in an idiom, as opposed to arguments introduced by a Low Applicative, which can indeed be a fixed part in an idiom.

I have argued that the present discovery about datives and the 'No Agent Idioms' generalization are instantiations of the same general constraint, namely that idioms are bound to the VP phase. By exploiting W&M's minimalist theory of argument structure, I have further proposed that the boundary for the VP phase is the first argument-introducing head merged on top of the vP; this is precisely the head that closes off the extended projection of the verb. Under such an account, the phasal status of Voice and High Applicative is formally unified in a straightforward way. The ban on fixed agents and high applied arguments in idioms is thus a welcome empirical fact.

If the present proposal has any merit, it suggests that constraints on idiom size follow from design issues motivated by extralinguistic factors, namely computational efficiency. This is theoretically felicitous from a minimalist/bilingualistic perspective. Also, rather than resorting to semantic characterizations like 'No Agent Idioms' and similar constraints, the limits to idiomacity can most naturally be accounted for on structural terms, a fact that underlines the primacy of syntax over semantics. All in all, the view supported in this work has been that idiomacity arises through the interaction of lexical items in a local context, the phase, and not as a result of direct semantic composition.

The goal and scope of this short paper have been modest, and many questions are left open for future research, one of them being whether the predictions made by the proposed hypothesis are accurate from a broader cross-linguistic perspective. Moreover, the present study has only focused on the behavior of high applied arguments in unaccusative idioms, but it remains to be seen what happens in cases in which both High Applicative and Voice are projected in the structure: which of the two is the phase head in this case? To answer this, a thorough study of transitive idioms involving a high applied argument is required.

## References

- Bruening, Benjamin. (2010). Ditransitive asymmetries and a theory of idiom formation. *Linguistic Inquiry* 41: 519–562.
- Bruening, Benjamin. (2020). Idioms, collocations, and structure. *Natural Language and Linguistic Theory* 38: 365–424. <https://doi.org/10.1007/s11049-019-09451-0>
- Burzio, Luigi. (1986). *Italian Syntax*. Dordrecht: Reidel.

- Chomsky, Noam. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. (2000). Minimalist inquiries. In R. Martin, D. Michaels, J. Uriagereka & S.-J. Keyser (eds.): *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. (2001). Derivation by phase. In M. Kenstowicz (ed.): *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. (2005). Three Factors in Language Design. *Linguistic Inquiry* 36(1), 1–22.
- Chomsky, N. (2007). Approaching UG from Below. Interfaces + Recursion = Language? In U. Sauerland & H.-M. Gärtner (eds.): *Chomsky's Minimalism and the View from Syntax-Semantics*, 1–29. Berlin: Mouton de Gruyter.
- Cuervo, María Cristina. (2003). Datives at large. PhD dissertation, Massachusetts Institute of Technology.
- Cuervo, María Cristina. (2020). Datives as applicatives. In A. Pineda & J. Mateu (eds.): *Dative constructions in Romance and beyond (Open Generative Syntax 7)*, 1–39. Berlin: Language Science Press. DOI: 10.5281/zenodo.3744254
- den Dikken, Marcel. (2023). High and Low Applicatives of Unaccusatives: Dependent Case and the Phase. *Linguistic Inquiry* 54(3): 479–503.
- Halle, Morris, & Alec Marantz. (1993). Distributed Morphology and the pieces of inflection. In K. L. Hale & S. J. Keyser (eds.): *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 111–76. Cambridge, MA: MIT Press.
- Harley, Heidi. (2014a). On the identity of roots. *Theoretical Linguistics* 40.4: 225–276.
- Harley, Heidi. (2014b). Reply to commentaries: “On the identity of roots.” *Theoretical Linguistics* 40.4: 447–474.
- Harley, Heidi, & Megan Schildmier Stone. (2013). The ‘No Agent Idioms’ Hypothesis. In R. Folli, C. Sevdali & R. Truswell (eds.): *Syntax and its limits*, 251–273. Oxford: OUP.
- Kratzer, Angelika. (1996). Severing the external argument from its verb. In J. Rooryck & L. Zaring (eds.): *Phrase Structure and the Lexicon*, 109–137. Dordrecht: Kluwer.
- Marantz, Alec. (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
- Marantz, Alec. (1997). No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2): 201–225.
- Mateu, Jaume. (2023). Idiomatic fake object resultatives revisited: The role of syntactic argument structure in their aspectual composition. Paper presented at the *International Workshop on Maximalization Strategies in the Event Domain*, University of Debrecen, Hungary, April 17-19.
- McGinnis, Martha. (2001). Phases and the syntax of applicatives. *North East Linguistics Society*: Vol. 31: Iss. 2, Article 7.
- McGinnis, Martha (2002). On the systematic aspect of idioms. *Linguistic Inquiry* 33: 665–672.
- McGinnis, Martha. (2005). Painting the wall red for a few hours: A reply to Glasbey (2003). *Snippets* 10. <http://www.ledonline/snippets/>
- Perlmutter, David. (1978). Impersonal passives and the unaccusative hypothesis. *Berkeley Linguistics Society* 4: 157–189.
- Pineda, Anna. (2013). Double object constructions in Spanish (and Catalan) revisited. In S. Baaui, F. Drijkoningen, L. Meroni & M. Pinto (eds.): *Romance Languages and Linguistic Theory, Selected papers from ‘Going Romance’ Utrecht 2011*, 193–216. Amsterdam: John Benjamins.
- Pylkkänen, Liina. (2002). Introducing arguments. PhD dissertation, Massachusetts Institute of Technology.
- Pylkkänen, Liina. (2008). *Introducing arguments*. Cambridge, MA: MIT press.
- Svenonius, Peter. (2005). Extending the extension condition to discontinuous idioms. *Linguistic Variation Yearbook* 5(1): 227–263.
- Wood, Jim, & Alec Marantz. (2017). The interpretation of external arguments. In R. D’Alessandro, I. Franco & A. J. Gallego (eds.): *The verbal domain*, 255–278. Oxford: OUP.

# Proceedings of the 42nd West Coast Conference on Formal Linguistics

edited by Shweta Akolkar,  
Amber Galvano, Akil Ismael,  
Kang Franco Liu, and Line Mikkelsen

Cascadilla Proceedings Project   Somerville, MA   2025

## Copyright information

Proceedings of the 42nd West Coast Conference on Formal Linguistics  
© 2025 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-484-3 hardback

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the printed edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Merino Hernández, Enrique. 2025. (Im)possible Idioms in a Minimalist Theory of Argument Structure: Insights from Spanish. In *Proceedings of the 42nd West Coast Conference on Formal Linguistics*, ed. Shweta Akolkar et al., 250-257. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #3828.