On semantic plurality within the Spanish DP and *ad sensum* agreement:
A combinatorial variability approach*

Timothy Gupton (University of Georgia) and
Chad Howe (University of Georgia)

Abstract

Spanish *ad sensum* agreement involves a semantically-plural, but grammatically-singular subject, of the form [NP₁ [*de* NP₂]] 'NP₁ of NP₂', appearing with a plural verb. Apparent agreement with a structurally-lower NP₂ is problematic for syntactic accounts dependent on c-command (e.g. Reinhart 1976). Previous accounts identify 'head N' via stipulation and cannot predict variable agreement.

We propose a uniform account of collective agreement and *ad sensum*/attraction phenomena in Spanish. Our proposal predicts variable agreement when semantic and grammatical number clash, triggering a *Count* mechanism that calculates phi-proportions, an index that informs variation in subsequent subject-verb agreement, similar to Adger (2006).

(98 words)

Key words: variable agreement, semantic features, syntax, *ad sensum*, DP, Spanish

1. Introduction

In this article, we detail a proposal to account for variable subject-verb agreement phenomena in complex, typically (pseudo)partitive noun phrases in Spanish of the type [NP₁ [*de* NP₂]] 'NP₁ of NP₂', as in (1) and (2), as found in the *Corpus del Español* (Davies, 2016).

- (1) *En España, por ejemplo, **ha vuelto la mayoría de los científicos**, aunque haya sido después de varias décadas.*

‘In Spain, for example, **the majority (sg.) of the scientists (pl.) have (sg.)** returned, even after several decades.’

- (2) *Pero **la mayoría de los músicos** de las orquestas españolas **son** extranjeros.*

‘But the **majority (sg.) of the orchestral musicians (pl.) are (pl.)** foreigners.’

In (1) the postverbal DP *la mayoría de los científicos* is the subject of the singular verb form *ha vuelto*, while in (2), the preverbal DP *la mayoría de los músicos* is the subject of the plural verb form *son*.¹ These two examples share a structure with two nominal elements that differ in grammatical number: singular in *la mayoría* and plural in *los científicos/los músicos*. Consider the DP *la mayoría de los músicos*:

(3)	<i>la</i>	<i>mayoría</i>	<i>de</i>	<i>los</i>	<i>músicos</i>
	def.f.sg	majority.f.sg	of	def.m.pl	musicians.m.pl
	'the majority of the musicians'				

In (2), the structurally lower, yet linearly closer NP₂ *músicos* appears to agree with the verb, resulting in what is often referred to as *ad sensum* agreement (e.g. Brucart, 1997).²

Here, we assume a syntax based on Chomsky (1995 *et sequens*), which assumes that the clausal structure is built up by recursive application of pair-wise Merge of lexical

and functional items from the Numeration, which is a subset of items selected from the Lexical Array for a particular utterance. We assume the advances of Abney's (1987) DP Hypothesis, as well as an extension of the NP to include functional projections, such as nP, the number-related NumP (see e.g. Ritter 1991; Picallo 2008), and KP providing Case (as in e.g. Lamontagne & Travis 1987) for partitives, thus yielding the clausal hierarchy in (4).³

(4) KP > DP > NumP > nP

Following the tenets of a feature matching and valuation approach as in Pesetsky & Torrego (2007), we assume that determiners (Ds) can be valued just once with respect to number, while nouns are endowed with grammatical number features [g] as well as semantic number [s] features.

For Corbett (1994), semantics and syntax are both actively involved in agreement processes. Chierchia (1998) suggests that mass nouns come from the lexicon already pluralized—even grammatically singular ones. Independent evidence in favor of such a possibility comes from Landau's (2000: 61) examination of control phenomena, in particular, partial control (5).

(5) John_i wanted/decided/refused [PRO_{i+} to meet at six]. (Landau, 2000: 44)

In (5), PRO_{i+} necessarily refers to John in addition to other parties. Landau suggests that semantic number is intrinsic to DPs, while syntactic number is arbitrary. He gives the example of *committee*, which is semantically plural, but syntactically singular. He proposes then that, in a similar fashion, PRO exists in the lexicon in two distinct 'flavors', a semantically-plural one [+SP], which can appear with partial-control (PC) predicates like (5, 6b) and a semantically-singular one [-SP] that appears with exhaustive control (EC) predicates, thus ruling out (6a).

- (6) a. *John told Mary that he managed to meet at 6 today. (EC)
 b. John told Mary that he preferred to meet at 6 today. (PC)

We interpret this as independent justification for the idea that number may be encoded in at least two separate areas within the DP. Consider now a model of the DP from (3):

As we will see presently, variable agreement is only indirectly influenced by the structurally-lower plural DP within the partitive KP *de los músicos* 'of the musicians'. At the point in the derivation when NumP merges with D, singular number must be valued on D, but the feature mismatch between g-number and s-number in the structurally-highest noun *mayoría* in NumP poses a valuation dilemma threatening crash. To save the derivation, g-number values are copied onto DP/D, subsequently triggering the operation *Count* at the DP phase edge.

(8) *Count*

Calculate a proportion related to phi-feature values when more than one feature value is present for the valuation of an uninterpretable feature. This proportion provides variable value instructions for subsequent agreement purposes.

We propose that, when a mismatch arises between grammatical and semantic number, the computation executes this operation, which tallies up grammatical and semantic number features. In this instance, there are three singular features and seven plural features, resulting in $\{3/10\}$. This proportion is a sort of diacritic that informs PF about variable agreement possibilities, similar to the combinatorial variability proposal of David Adger and colleagues (e.g. Adger, 2006; Adger & Smith, 2010).⁴ The proposed phase-edge calculus is not inconsistent with Chomsky's (1995: 230) comments regarding PF-mapping, "...mapping to PF has completely different properties and eliminates features in

ways not permitted in the $N \rightarrow \lambda$ computation—in particular, it eliminates formal and semantic features." Our proposal can straightforwardly explain variable agreement phenomena between a grammatically singular, yet semantically plural subject and either a singular or plural verb, but without having to stipulate a head noun or dispense with standard structural assumptions regarding agreement and c-command. In our proposal, the hierarchically lower DP within the (pseudo)partitive structure does not agree with the verb, but its features have an indirect influence on subject-verb agreement via *Count*.

In the following section, we examine previous accounts of *ad sensum* phenomena as well as their limitations. In section three, we return to the model we propose and examine some of its additional applications.

2. Previous analyses

According to Pérez-Jiménez & Brucart (2017), *ad sensum* agreement phenomena are attested in a wide number of languages, such as Portuguese, Catalan, Romanian, Russian, French, Arabic, German, and Hebrew.⁵ Examples of the *ad sensum* phenomenon in Spanish and its multiple subject-verb agreement possibilities appear in Gili Gaya (1980: 31):

- (9) a. *Un tropel de visitantes se ha /se han*
 indef.m.sg horde.m.sg of visitors.m.pl **SE aux.3.sg/SE aux.3.pl**
 reunido a la puerta.
 gather.ptcp at the door

'A horde of visitors has gathered at the door'

b. <i>La</i>	<i>mitad</i>	<i>de los</i>	<i>náufragos</i>
def.f.sg	half.f.sg	of def.m.pl	castaways.m.pl
<i>se salvaron</i>		/	<i>se salvó.</i>
SE save.PST.3.pl		/	SE save.PST.3.sg

'Half of the castaways were saved'

The diachronic literature has dedicated a considerable amount of attention to this sort of structure (e.g. Aarts 1998; Brems 2010; Traugott 2008a, 2008b), positing that constructions like English *a lot of NP* and *a sort of NP*, undergo a series of developments, beginning primarily as partitive constructions and then acquiring more general quantificational meanings. In the case of English structures like *a lot of*, these items have grammaticalized into ‘binominal quantifiers’, collocating a range of NP₂ types (e.g., count and mass) and extending beyond the nominal domain to modify vPs, as in ‘I go to the library *a lot*’. In Verveckken (2012: 436) the Spanish *un montón de libros* ‘a mountain of books’ is proposed to undergo a shift in literal meaning to a more abstract meaning equivalent to ‘a heap/a lot’. As part of this semantic change, she proposes that a structural reanalysis called ‘rebracketing’ takes place (10):

(10) [Det + QN_{head} + [de + N2] > [[Det + QN + de] + N2_{head}]

In Verveckken's proposal, QN refers to a 'quantifying noun' and the label 'head' designates the noun governing the entire DP. This proposal is that it seeks to capture the semantic and syntactic change in the *ad sensum* construction, while also capturing the fact that the NP₂ may ostensibly agree with the verb (11).

- (11) *Un centenar de soldados*
indef.m.sg hundred-some.m.sg of soldiers.m.pl
[detuvo / detuvieron] a los enemigos.
detain.PST.3.sg / detain.PST.3.pl DOM def.m.pl enemy
'a handful of soldiers stopped the enemy'

Although this sort of model attempts to explain how agreement may pass from the higher NP to the structurally lower NP, it does not explain why it sometimes does not happen. The proposal in (10) also runs into problems within a strictly generative, binary-branching syntactic model, as does Jackendoff's (1968) account, which similarly determines the 'head noun' by stipulation. In short, both of these models can arrive at the descriptive generalization that variable agreement occurs with *ad sensum*, but they lack explanatory power in that they cannot: 1) predict when *ad sensum* agreement will be impossible, or 2) predict when variation should occur.

Let us briefly examine the conundrum arising from attested *ad sensum* constructions *un tropel de visitantes* from (9a).

Assuming our simplified tree in (12), we see that the DP *un tropel* is hierarchically higher than the DP *visitantes* within the partitive KP. Upon merger of the structure in (12) with a finite T node, only the higher, singular DP *un tropel* will be in a c-command relationship with T and be able to agree with it. Therefore, according to these basic assumptions of generative syntactic theory, the lower, plural DP *visitantes* should never be able to agree with finite T, given that it is too deeply embedded in the hierarchical structure to be able to enter into a c-command relationship with T. Nevertheless, as we have seen the empirical facts are otherwise. The challenge then, is to find an alternate model of the syntax of partitive DPs that can account for when *ad sensum* agreement occurs, but also not generate any of the impossible instances of agreement mismatch.

As we can see, the standard Agree-based model in (12) cannot account for *ad sensum* data in any straightforward manner. What is more, it is incompatible with previous attempts to determine 'head noun' via stipulation. Unless we want to dismiss all such occurrences of mismatched agreement as performance errors—an undesirable prospect in our opinion—we need to consider improvements to the syntactic model. We believe that our model can explain not only *ad sensum* agreement phenomena, but also

collective nouns such as *gente* 'people' that also exhibit variable agreement behavior.

Before returning to our model, however, we examine the virtues and minor shortcomings of two additional proposals: Brucart (1997) and Pérez-Jiménez & Demonte (2017).

For Brucart (1997) there are two key characteristics that differentiate partitive constructions from pseudo partitives.⁶ The first is that partitives have a definiteness requirement within what Brucart calls the 'partitive coda', or the DP complement within the partitive structure, thus prohibiting (13a, cf. 13b-d).⁷

- (13) a. **Un grupo de* [_{DP} **unos** *manifestantes*]
 a group of D.indef.m.pl protestors
 'a group of some protestors'
- b. *Un grupo de* [_{DP} **cien** *manifestantes*]
 a group of hundred.m protestors
 'a group of a hundred protestors'
- c. *Un grupo de* [_{DP} **unos** **cien** *manifestantes*]
 a group of D.indef.m.pl hundred protestors
 'a group of approximately one hundred protestors'
- d. *Un grupo de* [_{DP} **unos** **pocos** *manifestantes*]
 a group of D.indef.m.pl few protestors
 'a group of a few protestors'

Despite their similar structures, pseudopartitives differ from partitives in that the coda DP (*niños* 'children' in (14)) must always be a bare nominal.

- (14) *Un grupo de niños*
 a group of children
 'a group of children'

Furthermore, true partitives allow extraposition of the partitive DP coda (15a, b), while pseudopartitives do not (15c, d):

- (15) a. *De los manifestantes, un grupo ~~de los manifestantes~~ se dirigieron*
 of the protestors a group of the protestors SE address.PST.3.pl
al gobierno civil.
 DOM-art.def.3.sg government civilian
 'of the protestors, a group addressed the civilian government'
- b. *De los manifestantes, el gobernador cree que la mayoría*
 of the protestors the governor think.PRS.3.sgthat the majority
~~de los manifestantes~~ eran estudiantes.
 of the protestors be.IMPV.3.pl students
 'of the protestors, the Governor thinks that the majority were students'
- c. **De manifestantes, un centenar ~~de manifestantes~~ se dirigieron*
 of protestors a hundred-some of protestors SE address.PST.3.pl

al *gobierno civil.*
 DOM-art.def.3.sg government civilian
 'of protestors, a hundred engaged with the civilian government'
 d. **De manifestantes, el gobernador cree que sólo un centenar*
 of protestors the governor think.PRS.3.sgthat only a hundred-some
~~*de manifestantes*~~ *eran estudiantes.*
 of protestors be.IMPFV.3.pl students
 'of the protestors the Governor thinks that only about a hundred or so were
 students'

This extraction asymmetry appears to be the prime motivator of Brucart's analysis of
 partitives and pseudopartitives, leading to an account that moves partitive coda DPs to an
 operator position adjoined to the DP at Logical Form (LF). Given that this is a semantic
 issue, he reasons that it should be resolved via covert movement at LF. However, as
 Pérez-Jiménez & Demonte (2017) correctly note, partitives as well as pseudopartitives
 trigger *ad sensum* agreement (16).

- (16) a. *La mayoría de los invitados llegó /*
 def.f.sg majority.f.sg of def.m.pl guests.m.pl **arrive.PST.3sg /**
llegaron en góndola.
arrivePST.3pl in gondola
 'The majority of the guests arrived via gondola.'

b. *Un puñado de soldados detuvo / detuvieron*
 def.m.sg handful.m.sg of soldiers.m.pl detain.PST.3sg / detain.PST.3.pl
a los enemigos.
 DOM def.m.pl enemy.m.pl
 'A handful of soldiers detained the enemy.'

In short, the fact that both structures trigger variable singular or plural agreement limits the explanatory power of Brucart's account. This is perhaps a welcome result, given the implications it would hold for this reconception of semantic module, which would find itself tasked with at least a subset of grammatical agreement calculations.

Stickney (2009) adopts the functional projection Measure Phrase (MP, Alexiadou, Haegeman & Stavrou 2007) for English pseudopartitives, proposing that the higher NP is an MP. While this would help to explain how a plural Number feature might 'percolate' upward from the coda NP to the head NP—a preferable alternative to determining head NP via stipulation—it leaves unresolved the issue of how singular agreement might happen, thus blocking percolation with pseudopartitives.

Pérez-Jiménez & Demonte's (2017) analysis does not depart from the basics of the generative understanding of agreement as related to c-command, as described above, such that the DP and the N head within agree with the verb. The innovative part of this model lies in its compositional approach to feature sharing and valuation: nouns and determiners are proposed to enter the syntactic derivation endowed with two sets of phi-features: c-features and *i*-features. C-features are concord features, which encode

morphosyntactic properties with relevance for the phonological form (PF) interface, and *i*-features encode semantic properties inherent in lexical items with relevance for the logical form (LF) interface.⁸ C-features include the formal, grammatical features Gender, Number, and Case, while the *i*-features include semantic properties with features for Gender, Number, and Person. In the typical instance of nouns, such as *niña* 'girl', the feature array in (17) is proposed:

- (17) *niña* C[G(f) N(sg) C()]
 i[G(f) N(sg) P(3)]

Nouns like *mayoría* 'majority' are different in that they have the option as appearing with their *i*-features valued, similar to *niña* in (17), or unvalued, as in (18).

- (18) *mayoría* C[G(f) N(sg) C()]
 i[G(f) N(sg) P(3)] / *i*[G() N() P()]

When unvalued, genitive-selecting nouns like *mayoría* get their *i*-features valued by the DP-compliment within the genitive KP. If that NP happens to be plural, those plural number feature values become the *i*-feature values of *mayoría* as well, resulting in the ability to agree with a plural verb form. Let us consider an example of how this works with (19).

- (19) *la mayoría de las mujeres*
 def.f.sg majority.f.sg of def.f.pl women.f.pl
 'the majority of women'

In (19), the N head *mayoría*, which begins with the c-features [G(f), N(sg), C()] and the as yet unvalued *i*-feature set $i[G(), N(), P()]$. This syntactic object then merges with the KP *de las mujeres* with the c-features [c[G(f) N(pl) C()] and the *i*-features [G(f) N(pl) P(3)]. Upon Agreeing, its *i*-feature values become those of *mayoría* (18').

- (18') *mayoría* c[G(f) N(sg) C()]
 $i[G(f) N(pl) P(3)]$

Consider when a singular DP appears within the KP, like *la mayoría del congreso* 'most of the conference'.

- (20) *congreso* c[G(m) N(sg) C()]
 $i[G(m) N(sg) P(3)]$

When (20) merges with *mayoría*, its *i*-feature N-value will become N(sg). This state of affairs should produce variable agreement for DPs like *la mayoría de las mujeres*, but never for DPs like *la mayoría del congreso*. It remains unclear, however, if the feature matrix in (20) for *congreso* is quite right, since it can also be understood as a plural

entity, similar to *mayoría*. If so, then it should have two possible *i*-feature sets (20'), similar to (18).

$$(20') \quad \text{congreso} \quad c[G(m) N(\text{sg}) C()]$$

$$i[G(m) N(\text{sg}) P(3)] \quad / \quad i[G() N() P()]$$

However, if no *c*-commanded complement appears with plural *N* *i*-feature values with which to value *i*, Pérez-Jiménez propose that plural number will be anaphorically recovered. It seems to us that features related to plural antecedents from preceding discourse must be available for at least some DP constructions. Otherwise, we would have no clear explanation for examples like (21) from Brucart (1997: 161, exs. 4a, b).⁹

- (21) a. **La mayoría es pariente.*
 def.f.sg majority.f.sg be.PRS.3.sg relative.m.sg
 'The majority is (a) relative.'
- b. *La mayoría son parientes.*
 def.f.sg majority.f.sg be.PRS.3.pl relative.m.pl
 'The majority are relatives.'

While this account fares well for examples with accompanying plural antecedents, it is unclear how semantically-plural *Ns* without antecedent constructions, such as *congreso* 'conference' or *gente* 'people' should ever have their *N* *i*-feature valued as plural. This

state of affairs would predict uniformly singular agreement for the grammatically singular DPs in *la mayoría de la gente* 'the majority/most of the people'. However, a cursory Google search reveals that this prediction does not bear out, as the examples in (22) attest:¹⁰

- (22) a. *De todas formas no me gustan para nada los debates, y **la mayoría de la gente opinan** igual que yo, solo sirven para reprocharse unos a otros cuatro tonter[i]as, y los temas principales ni tocarlos, es lo mismo que pasó en el último debate Rajoy-Rubalcaba.*

'All the same, I don't like these debates at all, and **the (f.sg) majority (f.sg) of the people (f.sg) think (pl.)** the same as I do, that they only serve to berate each other with four idiotic comments, and the main issues, forget it, it's the same that happened in the last Rojoy-Rubalcaba debate'

[<https://www.foro-ciudad.com/jaen/mancha-real/mensaje-9768105.html>, Mar. 12, 2012]

- b. *Veo que **la mayoría de la gente están muy ilusionada** con Griezmann, yo no, pero ya he expresado mi opinión y debatido sobre el tema, así que a partir de ahora no negativizaré más sobre Griezmann porque no tiene sentido ya seguir sobre el tema, y además tampoco quiero si **la mayoría está ilusionada**, cortar más el rollo.*

'I see that **the (f.sg) majority (f.sg) of the (f.sg) people (f.sg) are really dazzled (f.sg)** by Griezmann, not me, but I have expressed my opinion and debated on

the issue, so from now on, I won't go on about Griezmann because it doesn't make sense to keep talking about it, and besides, I don't want—if **the (f.sg) majority (f.sg) is (sg) dazzled (f.sg.)**—to be a wet blanket'

[<https://indiosrojiblancos.foroactivo.com/t4428p90-atleti-al-dia>, Sept. 6, 2021]

c. Pero *la mayoría de la gente no tienen* alternativas de ingreso.

'But **the (f.sg) majority (f.sg) of the (f.sg) people (f.sg) don't have (pl)** alternative income.'

[<https://www.elespectador.com/economia/economia-y-crisis-propuestas-para-enfrentar-la-pandemia-article-912345/>, Apr. 1, 2020]

Despite its elegance, the model proposed in Pérez-Jiménez & Demonte (2017) is not capable of predicting or explaining these instances of plural agreement. This is why we suggest that plural agreement has to come from another source. We return to our model in the following section, paying particular attention to its additional explanatory capabilities.

3. Extensions of the current proposal

Descriptively, we have seen that the nouns of interest are grammatically singular, yet semantically plural. That Number as a phi-feature should have a notional, semantic variant as well as a formal grammatical variant is hardly surprising if we consider similar features like gender and person. In a cognitive examination of agreement phenomena, Kibrik (2019: 74-75) correctly notes that, in Spanish (and Polish), third person deferential forms (i.e. *usted, ustedes*) are examples of *conventional* (i.e. grammatical) rather than

notional agreement. Similar examples, in which a grammatically third-person pronoun is used for second-person deferential reference, are abundant in a number of European languages. Kibrik also notes that gender in Russian is largely notional (i.e. semantic) with human referents. Spanish would appear to be similar, with the important exception of nouns with epicene agreement like *la víctima* 'the victim', which is grammatically feminine regardless of the gender of the human referent. We see this as additional motivation for a model incorporating semantic as well as grammatical phi-features, and have followed Pérez-Jiménez & Demonte (2017) in conceiving plural number as a separately-encoded semantic feature. However, as we have seen in their model, an extra set of semantic features alone does not make valuable predictions about variable agreement.

Adger (2006) examines variable subject-verb agreement in Buckie Scottish English (e.g. *we were* vs. *we was*) adopting a fairly standard Minimalist feature-checking system as in Chomsky (1995). Importantly, this variation is not associated with a variable meaning, nor is it influenced by social factors. In his account, lexical items may be underspecified for uninterpretable phi-features, and this is what can generate the patterns of variability and categoricity. He proposes that the following minimal phi-features are associated with a verb-form exponent (23):

- (23) a. [*usingular*: +] was
 b. [*usingular*: -] were
 c. [*uparticipant*: +] was

- d. [*u*author: -] were
- e. [*u*author: +] was

For the subject pronoun *you* [plural], the phi-feature set (24) results and, as we can see, this particular phi-combination has two paths to generating *was* from (23).

- (24) [singular: +, participant: +, author: -] (a) was, (c) was, (d) were

Following his proposal, the relevant feature matrices predict frequencies of variation as they are expected to occur in spoken corpora gathered in a typical sociolinguistic interview. In this case, two out of three outcomes result in *was*, so the predicted frequency should be 67%. In the corpus examined, the attested frequency of second-person singular *was* was 69%, which is almost identical to the predicted outcome. Although the sample size was small, combinatorial variability is a potentially promising heuristic for capturing and predicting agreement variation.

In the vein of Pérez-Jiménez & Demonte's (2017) proposal, we have posited a grammatical number feature *g* and a semantic feature *s* to capture the fact that a noun can at once be semantically plural and grammatically singular.¹¹ Why should a grammatically-singular DP result in plural verb agreement? The semantic plurality of collectives and (pseudo)partitives inherently allows for all members of a group act together, or separately in a variety of permutations, thus resulting in a distributive reading. This latter possibility allows for iterated instances of the predicate in question,

and this is precisely what implies plurality. Now let us re-examine a pseudopartitive DP in (25).

- (25) *un grupo de senadores*
 Art.m.sg.indef group of senator.pl
 'a group of senators'

In (26), we center our attention to Number (N) features in order to streamline the presentation and illustrate how this model avoids overgeneration.

- (26) a. $un_{[\phi:sg]} grupo_{[g:sg, s:pl]} de senadores_{[g:pl, s:pl]}$

 b. $*un_{[\phi:sg]} grupo_{[g:sg, s:pl]} de senador_{[g:sg, s:sg]}$

 c. $*un_{[\phi:sg]} grupos_{[g:pl, s:pl]} de senadores_{[g:pl, s:pl]}$

 d. $*unos_{[\phi:pl]} grupo_{[g:sg, s:pl]} de senadores_{[g:pl, s:pl]}$

In (26a), we see a standard example of *ad sensum*, where the phi-features on D are valued as singular. Note that s-feature matching should hold between NP₁ and NP₂; in (26b), the lack of an s-number match between NP₁ and NP₂ predicts that this result should be

impossible, true to fact. In (26c), we see that phi-features on D cannot receive a singular value because only plural is present in the phi-set of *grupos*, thus correctly ruling out this possibility. In (26d), we see that a clash of s-number and g-number cannot value D as plural.¹² We will return to this apparent problem for the current proposal shortly. First note that, in (27, 28), the same processes are at play as in (26a), when NP₁ and NP₂ are both semantically plural, but grammatically singular:

- (27) *un* *grupo* *de* *mi* *pueblo*
 indef.m.sg group.m.sg from my hometown.m.sg
 'a group from my hometown'

- (28) [DP [D *un* [ϕ :sg]] [NumP *grupo* [g:sg, s:pl] ... [KP *de* [DP *mi* [ϕ :sg] [NumP *pueblo* [g:sg, s:pl]]]
-

In (27), *pueblo* ‘hometown’ is conceptually (i.e. semantically) plural in that it is a collection of people and dwellings. Once again, we have a case in which s-number and g-number match between NPs. The plural s-feature of the highest NumP matches the s-plurality of the lower NumP. Also, in each NP the same type of g-number and s-number conflicts are present, thus setting the stage for variable agreement phenomena. This sort of structure, with a series of grammatically singular collective nouns, also creates an environment for *ad sensum* agreement (29).

- (29) ...*la mayor[i]a del estudiantado quieren*
 def.f.sg majority.f.sg of-def.m.sg student.body want.PRS.3pl
la suspensi[ó]n total de las clases hasta enero.
 the suspension total of def.f.pl classes until January
 'the majority of the student body want classes suspended
 [<https://diarioeltiempo.wordpress.com/2008/12/03/policias-y-vandalos-midieron-fuerzas-en-batalla-campal/>]

Examples like this one, from a 2008 Venezuelan discussion board post, pose a problem for Pérez-Jiménez & Demonte's analysis, which is dependent on grammatically-plural coda antecedents.

Two important problems in the current discussion remain that are not resolved by a simple addition of semantic features to the valuation mechanism of the model. The first is that it is unclear how a plural s-feature of the type we have proposed here should be able to trigger plural agreement with the verb if the DP is valued as singular (26a). The second problem is a related one, since as we have seen in (26d), in the event of a g-feature and s-feature number mismatch, plural number valuation on the DP results in ungrammaticality. Nevertheless, it seems that this is the step necessary in order to overcome the locality problem in (12) and allow the DP to participate in plural agreement with the verb. Consider the derivation of (30) up to the final merge operation within the DP:

At the point in which D merges with NumP, D requires phi-valuation, which importantly includes number. A phi-valuation problem at the phase edge arises owing to the non-matching semantic and grammatical features in NumP.¹³ The mismatch of g- and s-number initially forces a valuation of number on D identical to the g-number value on NumP in order to save the derivation. If this Last Resort operation did not apply, we would generate impossible, grammatically-mismatched DP combinations like (26d). In the partial Numeration in (30), Count tallies six singular features and four plural features out of a total possible ten Number features in the DP. We propose that it is this proportion $\{6/10\}$, resulting from *Count*, that allows and influences variable production, and remains as a special PF-index, as in (31).

This predicts the following probabilistic distribution of variable structures in the performance systems: roughly 60% of structures should have singular verb agreement and 40% should have plural agreement. As we have proposed, the *Count* process works in the spirit of Adger's (2006) *combinatorial variability* to "predict (rather than capture) frequencies of variants".¹⁴

What about collective nouns like *la gente* '(the) people'? We propose that these will be endowed with the phi-set [g:sg, s:pl], which also clashes and triggers the *Count* operation, valuing g-number on D as [g:sg] in the act. This leaves us with two singular features and one plural feature, thus predicting approximately 67% singular and 33% plural. Furthermore, partitive and pseudopartitive DPs with a clash at phase edge with a greater percentage of plural s-features in examples like *la mayoría de la gente de mi pueblo* 'the majority of the people from my hometown', which contains three grammatically singular, but semantically plural Ns, should result in greater proportions of plural agreement. Correspondingly, those with fewer plural s-features should result in lesser proportions of plural agreement. Future corpus studies will confirm or disconfirm these predictions.

4. Conclusions

There is always a tension present between descriptive adequacy and explanatory adequacy in formal syntactic analysis. When we are confronted with empirical data that challenge a proposed theoretical model, we must either modify the model to accommodate the data, or arrive at an informed decision as to why we should disregard the data and the insight that they might otherwise offer regarding modifications to the model. We feel that it is time that a generative syntactic model should be proposed that can make strides toward accounting for variable agreement data in *ad sensum* structures in languages like Spanish, thus capturing descriptive generalizations while making accurate predictions about the structure in question.

In the preceding, we have focused on *ad sensum* agreement in Spanish. We have examined its conception in the diachronic literature and the notion of re-bracketing, which amounts to little more than stipulating what the "head noun" (relevant for agreement) should be and restructuring the syntax around it, a state of affairs that will overgenerate singular and plural agreement. We have considered Brucart (1997), who accounts for *ad sensum* structures in an LF process that is dependent on covert movement. Unfortunately, this operation cannot predict the widely attested variable subject-verb agreement. We have also noted that the implications are concerning, since some portion of responsibility for grammatical agreement processes would then reside in the semantic module via LF. Pérez-Jiménez & Demonte (2017) attempts to account for the phenomena with concord (c) and index (i) phi-features. As we have shown, their model's dependence on plural Number valuation from the coda KP, however, limits its

coverage and does not provide an explanation for grammatically-singular but semantically-plural coda KPs. It also cannot account for the variability of subject-verb agreement.

Our proposal bears passing similarities to Pérez-Jiménez & Demonte (2017) in that we use two sets of phi-features: grammatical number (*g-number*) and semantic number (*s-number*). The co-existence of these features within an N is not a novel idea, and we have noted that it has independent motivation as related to partial control structures. The calculus that we propose is capable of generating the possible instances of of DPs that participate in *ad sensum* while ruling out the impossible ones. Unlike Pérez-Jiménez & Demonte (2017), however, our model can account for grammatically singular, but semantically-plural coda KPs. It can also account for collective nouns that lack an antecedent.

Although our account is inspired by Adger's (2006) combinatorial variability, we do not take an underspecification approach in our particular calculus, but we do not think it is necessarily incompatible.¹⁵ Future research will determine the degree to which underspecification is required for a feature-based account of variable agreement. This more nuanced approach to phi-features that we pursue via s-agreement and g-agreement matrices is a promising avenue for explaining apparently exceptional phenomena, however it is as yet unclear to what extent *Count* will be necessary to resolve clashes in other functional levels of the DP, or if this operation is only ever found at the phase edge.

A related concern for the DP is how diachronic change takes place and what forces lead to the near-categorical status of a particular expression, meaning that g-features

as well as s-features would become uniformly, or nearly uniformly, plural. For example, it seems to us that the English expression 'a lot of + NP' in *a lot of students* **attends/attend night classes*, is not possible with singular agreement, thus suggesting that its phi-matrix has become [g:pl, s:pl], similar to quantifiers like *many*. It is currently unclear to us if this is part of an organic historical change, or if prescriptive forces play a role in accelerating this sort of outcome. The diachronic syntax of partitives and pseudopartitives is taken up by Rutkowski (2007), working within the formal grammaticalization assumptions of Roberts & Roussou (1999). For the moment, however, we leave a deeper examination of historical change related to this structure for future research.

References

- Aarts, Bas. 1998. Binominal Noun Phrases in English. *Transactions of the Philological Society* 96(1). 117–158.
- Abney, Steven P. 1987. *The English Noun Phrase in its Sentential Aspect*. Cambridge, MA: MIT dissertation.
- Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42. 503-530.
- Adger, David and Jennifer Smith. 2010. Variation in agreement: A lexical feature-based approach. *Lingua* 120. 1109-1134.
- Alexiadou, Artemis, Liliane Haegeman & Melita Stavrou. 2007. *Noun Phrase in the Generative Perspective*. Berlin: DeGruyter.
- Bernstein, Judy. 1993. Word markers in null nominal constructions. *Probus* 5. 5-38.

- Bock, Kathryn & Cynthia A. Miller. 1991. Broken agreement. *Cognitive Psychology* 23. 45-93.
- Brems, Liesolette. 2010. Size noun constructions as collocationally constrained constructions: lexical and grammaticalized uses. *English Language and Linguistics* 14(1). 83–109.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Coon, Jessica and Keine, Stefan. 2021. Feature gluttony. *Linguistic Inquiry* 52(4). 655–710.
- Danon, Gabi. 2011. Agreement and DP-internal feature distribution. *Syntax* 14(4). 297–317.
- Davies, Mark. 2016. *Corpus del Español: Two billion words, 21 countries*.
<http://www.corpusdelespanol.org/web-dial/>
- Embick, David. 2015. *The morpheme*. Berlin: Mouton de Gruyter.
- Embick, David & Halle, Morris. 2005. On the status of *stems* in morphological theory. In Geerts, Twan, van Ginneken, Ivo, and Jacobs, Haïke. (eds), *Romance languages and linguistic theory 2003. Selected papers from 'Going Romance' 2003, Nijmegen, 20-22 November*, pp. 37-62. Amsterdam: John Benjamins.
- Embick, David & Noyer, Rolf. 2007. Distributed morphology and the syntax/morphology interface. In Ramchand, Gillian and Reiss, Charles (eds.), *The Oxford handbook of linguistic interfaces*, pp. 289-324, Oxford: Oxford University Press.
- Franck, J., U. Frauenfelder & L. Rizzi. 2007. A syntactic analysis of interference in subject-verb agreement. *MIT Working Papers in Linguistics* 53.173-190.

- Franck, Julie and Wagers, Matthew. 2020. Hierarchical structure and memory mechanisms in agreement attraction. *PLoS ONE* 15(5). 1-33.
- Gili Gaya, Samuel. 1980. *Curso superior de sintaxis española*. Barcelona: Biblograf.
- Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In Hale, Ken and Keyser, S. Jay (eds.), *The view from Building 20*, pp. 111-176, Cambridge: MIT Press.
- Harley, Heidi & Noyer, Rolf. 1999. Distributed morphology. *Glott International* 4(4). 3-9.
- Hartsuiker, Robert J., Schriefers, Herbert J., Bock, Kathryn, & Kikstra, Gerdien M. 2003. Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition*, 31(8), 1316–1326.
<https://doi.org/10.3758/BF03195814>
- Jackendoff, Ray. 1968. Quantifiers in English. *Foundations of Language* 4(4). 422-442.
- Kibrik, Andrej. 2019. Rethinking agreement: Cognition-to-form mapping. *Cognitive Linguistics* 30(1). 37–83.
- Lamontagne, Greg, & Travis, Lisa. 1987. *The syntax of adjacency*. In Crowhurst, Megan (ed.), *Proceedings of WCCFL*, pp. 173-186. Stanford: CSLI Publications.
- Landau, Idan. 2000. *Elements of control*. Dordrecht: Kluwer Academic Publishers.
- López, Luis. 2020. *Bilingual grammar*. Cambridge: Cambridge University Press.
- Marušić, Franc, Nevins, Andrew, & Badecker, William. 2015. The grammars of conjunction agreement in Slovenian. *Syntax* 18(1). 39-77.

- Pérez-Jiménez, Isabel & Violeta Demonte. 2017. Agreement and interpretation of partitive constructions in Spanish: The dual nature of nominal features. *Probus* 29(2). 355–395.
- Pesetsky, David & Torrego, Esther. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian and Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation. In honor of Joseph E. Emonds*, pp. 262-294. Amsterdam: John Benjamins.
- Picallo, M. Carme. 2008. Gender and number in Romance. *Lingue e Linguaggio* 7(1). 47–66.
- Reinhart, Tanya. 1976. The syntactic domain of anaphora. Doctoral dissertation, Massachusetts Institute of Technology.
- Ristić, Bojana, Molinaro, Nicola & Mancini, Simona. 2016. Agreement attraction in Serbian. Decomposing markedness. *The Mental Lexicon* 11(2).242–276.
- Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. In Susan Rothstein (ed.), *Syntax and Semantics vol. 25: Perspectives on phrase structure: Heads and Licensing*, 37–62. San Diego: Academic Press.
- Roberts, Ian, and Roussou, Anna. 1999. A formal approach to "grammaticalization". *Linguistics* 37. 1011-1041.
- Rutkowski, Paweł. 2007. The syntactic structure of grammaticalized partitives (pseudopartitives). *University of Pennsylvania Working Papers in Linguistics* 13(1). 337-350.

- Selkirk, Elisabeth. 1976. Some remarks on noun phrase structure. In Peter Culicover, Thomas Wasow and Adrian Akmajian (eds.), *Formal syntax*, pp. 285-316. New York: Academic Press.
- Slioussar, Natalia. 2018. Forms and features: The role of syncretism in number agreement attraction. *Journal of Memory and Language* 101. 51–63.
- Smith, Jennifer & David Adger. 2005. Variation and the minimalist program. In Leonie Cornips & Karen P. Corrigan (eds.), *Syntax and variation: Reconciling the biological and the social*, 149–178. Amsterdam: John Benjamins.
- Stickney, Helen. 2009. *The emergence of DP in the partitive structure*. Amherst: University of Massachusetts dissertation.
- Svenonius, Peter. 2004. On the edge. In Adger, David, de Cat, Cécile, and Tsoulas, George (eds.), *Peripheries: Syntactic Edges and their Effects*, pp. 259-287. Dordrecht: Kluwer Academic Publishing.
- Traugott, Elizabeth Closs. 2008a. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In Regine Eckardt, Gerhard Jäger & Tonjes Veenstra (eds.), *Variation, selection, development: Probing the evolutionary model of language change*, 219–250. Berlin: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 2008b. The grammaticalization of NP of NP patterns. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change*, 23–46. Berlin: Mouton de Gruyter.

- Verveckken, Katrien. 2012. Towards a constructional account of high and low frequency binominal quantifiers in Spanish. *Cognitive Linguistics* 23(2). 421–478.
- Wechsler, Stephen. 2011. Mixed agreement, the person feature, and the index/concord distinction. *Natural Language & Linguistic Theory* 29, 999–1031.
- Wechsler, Stephen & Zlatić, Larisa. 2000. A theory of agreement and its application to Serbo-Croatian. *Language* 76, 799–832.
- Wechsler, Stephen & Zlatić, Larisa. 2003. *The many faces of agreement*. Stanford, CA: CSLI Publications.

* Acknowledgements to be added following peer review.

¹ In the current proposal, we do not assume that subject position plays a role in variable agreement.

² Within the psycholinguistic literature, this phenomenon is often referred to as *attraction*. It has been examined in a number of languages including English (e.g. Bock & Miller 1991), French (e.g. Franck, Frauenfelder & Rizzi 2007; Franck & Wagers 2020), German (e.g. Hartsuiker, Schriefers, Bock & Kikstra 2003), Russian (e.g. Slioussar 2018), Slovenian (e.g. Marušić, Nevins & Badecker 2015), and Serbian (e.g. Ristić, Molinaro & Mancini 2016). It is worth noting, however, that the term *attraction* does not refer to a uniform phenomenon, and may be understood in multiple ways. While a more in-depth review of this literature is beyond the scope of this paper, in this study we limit our attention to grammatical instances of attraction related to this structure, and do not examine ungrammatical structures like (i).

(i) *The key to the cabinets are hard to reach.

³ Here we only assume NumP as a bare minimum with respect to the functional hierarchy of projections within the DP. With respect to Spanish, we do not assume that NumP is substantively different from the Word Marker Phrase (WMP) in Bernstein (1993). See Danon (2011) for further discussion of NumP crosslinguistically. Following rather standard practice in recent years following more widespread adoption of Distributed Morphology (e.g. Halle & Marantz 1993; Embick & Noyer 2007; Embick 2015), we assume a Root node ($\sqrt{}$) to form part of nP, which provides the root with its categorial identity. We abstract away from this level of detail for reasons of space.

⁴ This sort of approach is compatible with Late Insertion in Distributed Morphology (e.g. Harley & Noyer, 1999; Embick, 2015). Indices that distinguish homophones and diacritics to distinguish different classes, such as verb conjugations and noun declensions (as discussed in e.g. Embick & Halle, 2005) are proposed to exist on independent grounds. We thank **READER COLLEAGUE** for bringing this to our attention.

⁵ Slioussar (2018) in particular cautions that attraction phenomena in Russian differ from *ad sensum* phenomena in other languages. See also fn. 1.

⁶ See Selkirk (1976) for additional discussion related to partitives and pseudopartitives in English.

⁷ See also Jackendoff (1968) for similar structures and definiteness effects in English.

⁸ For more on this proposed feature matrix, couched within an HPSG framework, see Wechsler & Zlatić (2000, 2003), Wechsler (2011).

⁹ An open empirical question is what sort of variation is present among predicates. Example (i) (Brucart's (3b)) suggests that the predicate must play some role in this process.

-
- (i) La mayoría **consiguió/consiguieron** el triunfo.
the majority **achieve.PST.3.sg / achieve.PST.3.pl** the triumph
'Most were able to triumph'

This may mean a slightly different model of S-V agreement for copula. We leave this issue for future research. The agreement asymmetry may be related to what Coon & Keine (2021) call hierarchy effects of person and number features with copula in examples, similar to (ii-iv) in Spanish.

- (i) El inteligente *es/soy yo.
the intelligent be.PRS.3.sg / be.PRS.1.sg I
'I am the smart one'
- (ii) Los malos *son/sois vosotros
the bad be.PRS.2pl / be.PRS.3pl you.all
'You all are the bad guys'
- (iii) El problema *es/son ellos
the problem be.PRS.3.sg / be.PRS.3.pl they
'The problem is them'

Following their proposal, the most specified of these features should prevail in agreement terms within copula. We are grateful to **READER COLLEAGUE** for this observation and the examples.

¹⁰ These Google search results are from message boards about sports and politics (17a,b) from Spain and (17c) is from a Colombian newspaper. Here, in order to streamline the presentation, we combine glosses and idiomatic translation, highlighting the relevant agreement phenomena in bold.

¹¹ Feature matrices *g* and *s* do not possess any substantive differences that we are aware of from the *c*-features and *i*-features in Pérez-Jiménez & Demonte's analysis. The main difference lies in how we conceive of the system and under what circumstances it should 'count up' features

¹² Apparent counterexamples like *un/el microondas* 'a/the microwave' or *un/el lavaplatos* 'a/the dishwasher' are not hard to come by, but combinations like these are proposed to have a phi-matrix in NumP of [g:sg, s:sg].

¹³ As far as we can tell, nothing hinges on whether the phase edge is DP (e.g. Svenonius 2004) or KP, as assumed in López (2020), with DP inheriting features from KP. Regardless, we assume that the semantic number clash issue must be resolved prior to transfer of this phase. This process may also form part of the Conceptual-Intentional (C-I) interface between the syntax and the semantic module.

¹⁴ See also Smith & Adger (2006), Adger & Smith (2010).

¹⁵ We suspect that underspecification may be responsible for the choice of singular, the least-marked number value.