

Interpretable everywhere: Hybrid agreement in Brazilian Portuguese

Giuseppe Varaschin

Humboldt-Universität zu Berlin

 0000-0003-1446-2700

giuseppe.varaschin@hu-berlin.de

Antonio Machicao y Priemer

Humboldt-Universität zu Berlin

 ORCID ID: 0000-0001-7321-0795

machicao.y.priemer@hu-berlin.de



How to cite: Varaschin, Giuseppe & Antonio Machicao y Priemer. 2025. Interpretable everywhere: Hybrid agreement in Brazilian Portuguese. RLLT 26, eds. Pilar Barbosa, Cristina Flores, Esther Rinke, Eva-Maria Roessler. Special Issue of *Isogloss*. *Open Journal of Romance Linguistics* 11(7)/17 (1–45).

DOI: <https://doi.org/10.5565/rev/isogloss.576>

Abstract

The goal of this paper is to account for the variable agreement properties of the 1PL pronoun *a gente* in Brazilian Portuguese. We argue that previous approaches make wrong predictions about how different types of agreement with *a gente* work in different environments (Costa & Pereira 2013; Taylor 2009, i.a.). As an alternative, we argue for a dual-feature approach that incorporates the distinction between ϕ -features that are semantically interpretable (INDEX features) and those that determine phonological exponents (CONCORD features) (Smith 2021; Wechsler & Zlatić 2003). We propose that *a gente* has 1PL value for the former and a 3SG value for the latter and that both types of features can be targeted in non-anaphoric agreement (e.g. subject–verb agreement). The distinct usage preferences for different types of features in different environments follows from register considerations: structures where *a gente* is associated with overt 1PL exponents convey opposing social meanings. In order to capture this effect, we formulate a theory of social meaning composition based on formal models of honorification and conventional

implicatures (cf. Machicao y Priemer et al. 2025; McCready 2019; Potts 2007; Varaschin et al. 2024).

Keywords: agreement, concord-index distinction, pronouns, variation, ϕ -features, register, social meaning, conventional implicatures.

1. Introduction

Brazilian Portuguese (BP) employs two forms to express first-person plural (1.PL): the personal pronoun *nós*, with its clitic counterpart *nos*, and a newer form *a gente*, which is diachronically derived from the combination of the feminine definite determiner *a* with the collective singular noun *gente* ('people') (Alencar 2013; Carvalho et al. 2018; Costa & Pereira 2013; Lopes 1998, 1999, 2004; Marcotulio et al. 2013; Menuzzi 2000; Pereira 2003; Vianna & Lopes 2012; Zilles 2002, 2005, 2007).¹

A gente is truth-conditionally equivalent to *nós*. Like the latter, *a gente* denotes a plurality that includes the speaker, which can be seen in its acceptability with collective predicates in (1a). The main distinguishing property of *a gente* (in contrast to *nós* and 1.PL forms in general) is that, in spite of its semantics, it typically controls 3.SG agreement (Lopes 2004: 52-53). This is evident in cases of subject-verb agreement (1a) and anaphoric agreement (1b), where the less frequent agreement targets are marked with '?'. These less frequent variants are, to varying degrees, judged negatively (e.g. as not fully acceptable) by speakers of standard varieties (Freitag 2016; Menuzzi 2000, i.a.).

- (1) a. A gente {é / ?somos} uma família.
A GENTE be.3.SG be.1.PL a family
'We are a family.'
- b. A gente_i {se_i / ?nos_i} viu na TV.
A GENTE REFL REFL.1.PL saw.3.SG on-the TV
'We_i are saw ourselves on TV'

We show, using data from previous sociolinguistic studies and the BP section of *Corpus do Português* (Davies 2016), that previous approaches are too restrictive in their assumptions about *a gente*'s affordances to agreement (Carvalho et al. 2018; Costa & Pereira 2013; Marcotulio et al. 2013; Menuzzi 2000; Taylor 2009). This excessive restrictiveness stems from the attempt to rule out marked agreement patterns – like the ones signaled with '?' in (1) – by means of syntactic resources.

Though occurrences of *a gente* with 1.PL targets are rare, examples are consistently attested in corpora and sociolinguistic work (Brito & Sedrins 2017; Marcotulio et al. 2013; Taylor 2009; Varaschin 2021a). We take this to mean that such patterns are allowed by the grammar and that their markedness needs to be attributed to other causes. The purpose of this paper is, therefore, to account for *a gente*'s hybrid grammatical properties, as well as for its usage preferences. That is, we want to have a single grammatical

¹ Marcotulio et al. (2013: 141) show that, *pace* minor quantitative differences, the core facts about *a gente* reported here also hold for European Portuguese. In this paper, however, we focus on BP, where the behavior of *a gente* has been shown to be stable and entrenched across all dialects (Alencar 2013; Lopes 1999; Rubio & Gonçalves 2012; Seara 2000; Vianna & Lopes 2015; Zilles 2005, 2007).

machinery that derives not only the pool of possible agreement variants but also the fact that some of these variants are less frequent and less acceptable than others.

In order to do this, we opt for a richer architecture of features than the one assumed by previous accounts of *a gente* where (potentially conflicting) syntactic and semantic information of agreement controllers are simultaneously represented, serving as potential controllers for different types of agreement targets (Adamson 2025; Kathol 1999; Puškar 2017; Smith 2017, 2021; Wechsler & Zlatić 2003; Wurmbrand 2017). We adopt a dual-feature system based on Wechsler & Zlatić (2003), distinguishing between ϕ -features that constrain assignment functions (INDEX features, ϕ^I) and those that determine phonological exponents (CONCORD features, ϕ^C). The core of our proposal is that *a gente* is a pronoun whose INDEX value is 1.PL and whose CONCORD value is 3.SG. Both types of ϕ -features can control grammatical agreement, but only ϕ^I -features control anaphoric agreement, thus explaining the range of agreement variation and mismatches.²

The fact that some agreement patterns are less frequent and perceived to be less acceptable than others is explained in terms external to these core syntactic mechanisms. We argue that the markedness of *a gente+V_{1.PL}* and *a gente+nos* combinations follows from a compositional theory of social meanings, understood as conventional implicatures that place constraints on utterance contexts (Kaplan 1999; Machicao y Priemer et al. 2025; McCready 2019; Potts 2005, 2007; Varaschin et al. 2024, i.a.). Building on prior sociolinguistic results, we ascribe to *a gente* and to forms with 1.PL as ϕ^C -value conflicting social meanings – the former has a low-to-moderate degree of formality, while the latter have a very high degree of formality (Brustolin 2009; Freitag 2016; Lima 2015; Naro et al. 1999; Schwenter et al. 2022; Zilles 2005). Given the particular compositional principle we propose, this entails that the range of contexts where the combination between *a gente* and $\phi^C:\{1.PL\}$ forms can be felicitously used is very narrow.

In Section 2 we present the core data that theories dealing with *a gente* must grapple with, including its variable behavior with respect to different types of agreement (verbal, predicative and anaphoric). In Section 3, we review previous accounts of *a gente* and argue that they make wrong predictions. Section 4 sketches a dual-feature analysis for the grammar of *a gente* that overcomes these difficulties, predicting all and only the attested agreement patterns. Section 5 offers a register-based explanation for why the less frequent agreement patterns with *a gente* are often perceived as unacceptable.

2. Data

In this section, we outline basic facts that theories dealing with *a gente* need to account for. Our data come primarily from previous sociolinguistic studies and from the BP section of Corpus do Português: Web/Dialects (Davies 2016). The BP subset of this large corpus (1B words) is compiled by restricting results to webpages geolocated to Brazil by Google and by independently validating geographic provenance using lexical and mor-

² The particular implementation of the agreement mechanism underlying this dual-feature approach is not particularly relevant for our purposes. As long as the distinction between ϕ -features that appear as values of INDEX and CONCORD is acknowledged, any system of principles which ensures the token-identity of ϕ -features between agreement controllers and agreement targets suffices. Our proposal for the structure of *a gente* is, thus, compatible with both derivational and constraint-based theories of agreement.

phosyntactic cues associated with BP varieties. Throughout the article, corpus-extracted examples appear in quotation marks. Other examples in the paper are constructed, either by us or borrowed from previous literature (where noted), but every pattern illustrated is independently attested in the corpus and/or in quantitative sociolinguistic surveys cited throughout. We indicate sources of previously published acceptability judgments; in addition, all examples have been checked with at least two native speakers of southern BP.

Following Chomsky (1965: 9–10), we take it as given that native speakers do not have direct access to grammaticality. What we can observe empirically are acceptability judgments and patterns of spontaneous production, which reflect the interaction of multiple factors, only one of which is grammaticality. The relationship between observable measures and the notion of grammaticality is, thus, complex and theory-dependent. We assume that a descriptively adequate grammar should be consistent not only with the most frequent and productive variant of a linguistic variable, but also with its marked and sometimes less-than-fully acceptable realizations, provided these are consistently attested in intra-dialectal corpora and judged by speakers to be comparatively *more* acceptable than low-acceptability baselines that are unambiguously rejected (e.g. (16)). Ideally, grammatical theories should also incorporate results from sociolinguistic research identifying situational factors that influence the distribution of a variant, since such factors may constitute confounds for syntactic accounts of why a particular structure sounds degraded.

When a speaker judges a structure *S* to be unacceptable, this may simply reflect that they do not see themselves producing *S*. There are, however, numerous reasons, independent of grammaticality, why a speaker may avoid producing *S*. For instance, *S* may be associated with speakers of a particular social group which the speaker does not want to identify with (Eckert 2012). In other words, a structure may be syntactically licensed but unacceptable (to varying degrees) for reasons external to the grammar proper, such as pragmatic inappropriateness or negative social evaluation. This is precisely the type of argument we develop in Section 5 for marked agreement patterns with *a gente*.

In what follows, we start by establishing (contra Taylor 2009) the status of *a gente* as a 1PL pronoun. Then, we discuss its properties as an agreement controller with respect to verbal, adjectival and anaphoric targets, drawing on heavily on prior empirical work.

2.1. Pronominal status

Before looking at the behavior of *a gente* with respect to different types of agreement, it is important to establish that it is, in fact, a 1PL pronoun. We summarize some the main arguments for this which are, to a large extent, drawn from Menuzzi (2000: 203–208).

First, *a gente* has the range of interpretations typically associated with 1PL pronouns: it is a variable that refers to a specific plurality of individuals including the speaker and can also enter into anaphoric relations with the 1PL pronoun *nós*, as in (2).

(2) Só a gente_i viu uma cobra atrás {de nós_i / d-a gente_i}.
 only A GENTE saw.3SG a snake behind of us of-A GENTE
 ‘Only we_i saw a snake behind us_i.’

Second, *a gente* shares with plural pronouns the possibility of a distributive interpretation with mixed predicates, as in (3a). This contrasts with 3SG collective-denoting referential DPs like *a banda* (‘the band’), which only has a cumulative reading in (3b).

(3) a. A gente ganha R\$50000 por mês.
 A GENTE earn.3SG R\$50000 per month
 ‘We make R\$50000 a month.’ \rightsquigarrow each person in the group makes R\$50000

b. A banda ganha R\$50000 por mês.
 the.F.SG band earn.3SG R\$50000 per month
 ‘The band makes R\$50000 a month.’ $\not\rightsquigarrow$ each band member makes R\$50000

Third, *a gente*, like other pronouns (e.g. *eu*, *você*), can receive arbitrary/generic interpretations and enter into anaphoric relations with other impersonal subjects, as in (4) (examples adapted from Carvalho & Brito 2017: 62–63 and Menuzzi 1995: 156).

(4) a. A gente começa a comer, a gente não para mais.
 A GENTE begin.3SG to eat A GENTE not stop more
 ‘Once one starts eating, one doesn’t stop.’

b. Sempre se_i imagina que a gente_i pode escapar do perigo.
 always SE imagine.3SG that A GENTE can escape from.the danger
 ‘One always thinks that one can escape from danger.’

Fourth, like all other pronouns, including 1PL *nós*, and unlike regular referential DPs, *a gente* cannot be attributively modified by adjectives keeping its 1PL reading:³

(5) a. * [A gente feliz] começou a rir.
 A GENTE happy started.3SG to laugh

b. * [Nós felizes] começamos a rir.
 we happy started.1PL to laugh

c. [A galera feliz] começou a rir.
 the folks happy started.3SG to laugh
 ‘The happy folks started to laugh.’

Fifth, like other pronouns, *a gente* is subject to Principle B (6) and not subject to Principle C, cf. (2) above (i.e. it can be bound by its antecedent in a non-local domain):⁴

(6) * [A Carolina e eu]_i odiamos a gente_i.
 the Carolina and I hate.1PL A GENTE
 ‘[Carolina and I]_i hate us_i.’

Lastly, in some varieties of BP, *a gente* undergoes phonological reduction from the trisyllabic realization on the left in (7a) to one of the variants on the right of the arrow in (7b). The monosyllabic variants may sometimes lead to cliticization (Maia 2012; Ramos & Maia 2015; Zilles 2005) – a known prosodic property of pronouns in many languages (Cardinaletti & Starke 1999; Déchaine & Wiltschko 2002; Panagiotidis 2002).

³ As we show below, *a gente* and *nós* can appear with nouns in pronominal definite constructions like *a gente linguistas* (‘we linguists’). We assume these are different from attributive constructions like (5). Though some human nouns may double as human adjectives in BP, not all human adjectives can be nominalized. This is the case of *feliz* in (5), which cannot appear alone with determiners, as in **o feliz* (‘the happy’), or as a bare singular in generic sentences, as in **Feliz adora música* (‘Happy people love music’).

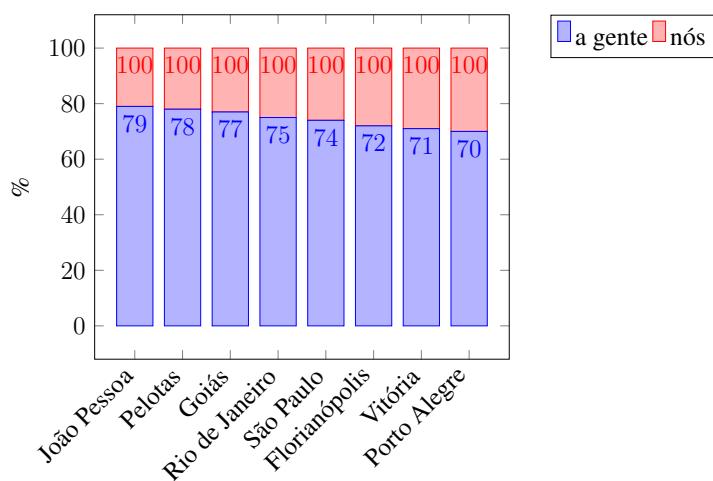
⁴ Varaschin (2021a,b) argues that personal pronouns in BP do not exhibit Principle B effects in the same way as English. However, BP does disfavor pronouns in contexts like (6a) (see also Carvalho 2021).

(7) a. [a.ʒẽ.tʃi] ⇒ b. [aẽ.tʃi], [ẽ.tʃi], [ʒẽ.tʃi], [ẽ.tʃi], [ʒẽ], [tʃi]

2.2. Variation

A gente is the most productive variant of 1PL across all varieties of contemporary BP. The prevalence of *a gente* vis-à-vis *nós* in spoken speech ranges between 79%–70% in virtually all regions of Brazil.⁵ Figure 1 shows that the distribution of *a gente* does not vary substantially across regional dialects. Other situational factors are stronger predictors for the occurrence of *a gente*, such as being of *younger age* (Maia 2009; Rubio & Gonçalves 2012; Seara 2000; Zilles 2005), *female* (Borges 2004; Mendonça 2012; Monteiro 1994; Seara 2000; Zilles 2005), and having a *lower level of education* (Lima 2015; Lopes 1998; Monteiro 1994; Omena 1996; Rubio & Gonçalves 2012; Seara 2000). Table 1 displays data concerning the social embedding of *a gente* in the state of Goiás.

Figure 1. First person plural variation in different varieties of BP



Attitude studies demonstrate that speakers also perceive *a gente* as less formal than *nós* (Brustolin 2009; Freitag 2016; Freitag et al. 2018; Freitas & Carvalho 2020; Lima 2015; Vianna 2006). Another indication of this, as Zilles (2007) points out, is the relative scarcity of *a gente* in literature, and its unclear status in dictionaries and normative grammars. This suggests that *a gente* not only correlates with informal registers in speech, but functions as *marker* of low formality, in the sense of Labov (1972).

Turning to its morphosyntactic behavior, *a gente* triggers two patterns of agreement with inflected verbs (verbal agreement, VA): 3SG and 1PL. We illustrate the attested patterns (without acceptability judgments) in the constructed minimal pairs in (8).⁶

⁵ The surveys whose results are summarized in Figure 1 are the following: Fernandes (2004) (*apud* Vianna & Lopes 2012: 98) (João Pessoa); Borges (2004: 129) (Pelotas); Mattos (2013: 108) (state of Goiás); Lopes (2004: 69) (state of Rio de Janeiro); Rubio & Gonçalves (2012: 1016) (countryside of São Paulo); Seara (2000: 181) (Florianópolis); Mendonça (2012: 4) (Vitória); and Zilles (2002: 302) (Porto Alegre). All of these studies deploy mixed samples (i.e. they were conducted with speakers of diverse age-groups, genders and educational levels). Variation exists both across and within groups.

⁶ In non-standard varieties, *nós* is also attested with what look like 3SG verbal targets, e.g. *nós corre* ‘we run’. This raises the question of whether 3SG agreement with *a gente* is in fact 3SG or some form of impoverished 1PL agreement (cf. Harley 2008; Nevins & Parrott 2010, i.a.).

Table 1. Social variables correlated *a gente*

Social variables	N/Total	Frequency	Weight
Age			
16-24	602/690	87%	.70
25-40	715/933	77%	.49
41-86	269/439	61%	.23
Education			
High School	703/812	87%	.69
College	883/1250	71%	.37
Gender			
Feminine	782/984	80%	.60
Masculine	804/1078	75%	.41
Total	1586/2062	77%	Input: .83

Source: Mattos 2013: 111-112

(8) a. A gente corre.
 A GENTE run.3SG
 b. A gente corremos.
 A GENTE run.1PL

In spite of (8a) being judged as more acceptable than (8b), both of these variants occur not only across groups (at different rates), but also within groups; see the reviews in Vianna (2006), Alencar (2013), Freitag (2016) and the references there. Since these studies control for external sociolinguistic factors (age, gender, education, etc.), the standard conclusion in variationist work is that the residual variability is intra-individual in nature (Fasold & Preston 2007; Labov 1969). This means that competent speakers know both variants, even when they avoid using (8b). We return to this point in the end of the section.

Some studies have also found 3PL VA controlled by *a gente*, as the corpus example in (9) shows (Alencar 2013; Marcotulio et al. 2013; Vianna 2006).⁷

(9) “[M]uito bom quando as pessoas que a gente amam se gostam.”
 very good when the people that A GENTE love.3PL REFL like.3PL
 ‘It is very good when the people that we love like each other.’

Table 2 displays the frequencies for all of the three VA patterns attested for *a gente* in written samples and written tests for speakers in Rio de Janeiro. These results show that the realization of plural agreement morphology on inflected verbs is much more prevalent

Though impoverishment is arguably active in the inflectional system of BP, this phenomenon is separate from the one we are dealing with here. First, *a gente* also controls predominantly 3SG verbal targets in European Portuguese (Pereira 2003), where there is no parallel impoverishment of verbal inflection going on. Second, the impoverishment rule that yields 3SG-like targets for *nós* has a very different social meaning from the type of rule that yields 3SG targets with *a gente*, given that the latter is quite standard and the former is not. If *a gente* were merely specified for 1PL and the variation in the form of agreement targets were explained in terms of impoverishment alone, we would have nothing to attribute this difference in sociolinguistic evaluation to.

⁷ The 3PL variant is even more strongly rejected by speakers of southern BP than 1PL variants are. Though quite marginal in BP, the 3PL pattern is the most common one in the variety of Portuguese spoken in the island of São Miguel in the Azores (Costa & Pereira 2013; Pereira 2003). We assume that the presence of variant is a matter of dialectal variation and is not grammatically licensed for most BP speakers.

in written registers, which suggests that the morphosyntactic PL feature itself may be seen as a marker of formality or more careful styles in BP (Alencar 2013; Brito & Sedrins 2017; Costa & Figueiredo Silva 2006; Zilles 2005). It has also been shown that 1PL VA with *a gente* is more common among older speakers (Mattos 2013: 107). Analyzing the diachronic data in Vianna (2006), Alencar (2013: 54) concludes that both 1PL and 3PL VA with *a gente* are decreasing in frequency over time, which could also indicate a social evaluation of plural marking as being more conservative and formal in contemporary BP.⁸

Table 2. VA with *a gente* in two modalities

	3SG	1PL	3PL
spoken samples	99% (1046/1054)	1% (8/1054)	0% (0/1054)
written tests	81% (334/411)	18% (73/411)	1% (4/411)

Source: Marcotulio et al. 2013: 132

Regarding predicative agreement (PA) in copular clauses (with adjectives, participles, etc.), *a gente* exhibits four possibilities: M.SG, M.PL, F.SG and F.PL. We present these possibilities with constructed minimal pairs, all of which replicate patterns found in Corpus do Português (Davies 2016) and prior quantitative studies (Vianna 2006):

(10) a. A gente está cansado.
A GENTE is tired.M.SG

b. A gente está cansados.
A GENTE is tired.M.PL

c. A gente está cansada.
A GENTE is tired.F.SG

d. A gente está cansadas.
A GENTE is tired.F.PL

As in the case of VA, these variants also vary in their distribution across written and spoken samples within groups of speakers, with SG being the most common form overall. Table 3 displays the distribution across samples for speakers from Rio de Janeiro.

Table 3. PA with *a gente* in two modalities

	M.SG	M.PL	F.SG	F.PL
spoken samples	90% (37/41)	0% (0/41)	10% (4/41)	0% (0/41)
written tests	70% (242/344)	14% (47/344)	13% (46/344)	3% (9/344)

Source: Marcotulio et al. 2013: 133

⁸ Lima (2015: 136) reports that 1PL verbal targets are more frequent in formal writing than in informal speech – even when the 1PL subject is *a gente*. Brustolin (2009: 214) confirmed this in a sociolinguistic perception study which shows that, regardless of whether *nós* or *a gente* is used as a subject, 1PL inflection on verbs is perceived as more appropriate in formal situations (see also Freitag 2016: 909).

In written samples, the preference for SG targets with PA is less strong. This is further indication that formality is a factor that contributes to a slight increase in morphosyntactic PL marking. Gender in PA with *a gente* is, in turn, determined by referential properties and not by arbitrary lexical specification (which would arguably be FEM – i.e. the inherent gender of the collective noun *gente*). This is another property that speaks in favor of *a gente*'s pronominal status, since among DPs, underspecification for gender is almost exclusive to pronouns. In their empirical study, Marcotulio et al. (2013: 135) found that FEMALE PA targets with *a gente* as a controller, as in (10c)–(10d), are only preferred for female-exclusive reference. In all other cases *male-exclusive, generic and mixed reference*, MALE PA has significantly higher frequency (over 80%).⁹

Combining these four options with the two patterns of VA (setting aside 3PL, which is rare and possibly dialectal), there is a total of eight possibilities of VA+PA combinations, all of which are attested (Vianna 2006). Crucially, empirical studies do not find evidence for one form of VA entirely blocking a mismatching form of PA: e.g. plural VA is possible with singular PA, and vice-versa, as the attested examples in (11) show. However, PA-VA mismatches are perceived as marked and degraded by most speakers.

(11) a. “[D]epois de 3 anos juntos a gente fomos indenizado pela
after of 3 years together A GENTE were.1PL compensated.M.SG by-the
urbeu [...]”
urbeu
'After 3 years together we were compensated by Urbeu [...]'
b. “[...] a gente é amigos mesmo.”
A GENTE is.3SG friend.M.PL really.
'We are really friends.'

We now turn to anaphoric agreement (AA) (Bresnan & Mchombo 1987): i.e. agreement between *a gente* and pronoun or reflexive which is bound by it. Unlike VA and PA, there is a scarcity of empirical studies on this issue. In non-local contexts, *a gente* can be anaphorically related to *a gente* or *nós* (and its corresponding clitic form *nos*).

(12) a. A gente_i acha que a Maria viu a gente_i.
A GENTE think.3SG that the Maria saw.3SG A GENTE
'We_i think that Maria saw us_i.'
b. A gente_i acha que a Maria nos_i odeia.
A GENTE think.3SG that the Maria CL.1PL hate.3SG
'We_i think that Maria hates us_i.'
c. A gente_i acha que a Maria gosta de nós_i.
A GENTE think.3SG that the Maria likes.3SG of us
'We_i think that Maria likes us_i'.

⁹ These same possibilities are attested for the standard 1PL form *nós*, but with different frequencies. In a study reported in Lopes (2004: 59), the most common PA pattern is MASC.PL (55%), followed by MASC.SG (34%), FEM.PL (%9) and FEM.SG (2%). These figures are similar to the PA data found for *a gente* in European Portuguese, where plural agreement with predicative adjectives is also the more frequent pattern (Costa & Pereira 2013; Marcotulio et al. 2013; Pereira 2003).

In local contexts, *a gente* preferentially binds the reflexive proclitic *se*. This is reported by Menuzzi 2000, Taylor (2009), Reuland (2011) as well as our BP informants:

(13) A gente_i se_i viu na TV.
 A GENTE REFL saw.3SG on-the TV
 ‘We_i saw ourselves_i on TV.’

Traditionally, *se* is taken to be a exponent for 3SG and 3PL anaphors. However, in contemporary BP, it also takes up other functions: it can have 1PL *nós*, 2SG *tu* and even 1SG *eu* as an antecedent, as Brito’s (2008: 489) examples in (14) show (Brito 2008; Menuzzi & Lobo 2016; Pereira 2007). For this reason we assume *se* is not 3SG, but a syncretic form whose semantic features are underspecified in both person and number.

(14) a. Eu_i se_i lavo todos os dias.
 I REFL wash.1SG all the days
 ‘I wash myself everyday’

b. Nós_i se_i lavamos todos os dias.
 we REFL wash.1PL all the days
 ‘We_i wash ourselves_i every day.’

c. Tu_i se_i lavas todos os dias.
 you REFL wash.1PL all the days
 ‘You_i wash yourself_i everyday.’

Crucially, *a gente* can also be a local antecedent for the 1PL reflexive clitic *nos* (Brito & Sedrins 2017; Carvalho et al. 2018; Taylor 2009; Varaschin 2021a), regardless of the form of the verb, as in the attested examples in (15) show.¹⁰

(15) a. “O amigo é um presente que a gente_i nos_i dá [...]”
 the friend is a gift that A GENTE REFL.1PL give.3SG
 ‘The friend is a gift that we_i give to ourselves_i.’

b. “É mais fácil a gente_i nos_i alegrarmos depois de passar o dia.”
 is more easy A GENTE REFL.1PL cheer.INF.1PL after of pass the day
 ‘It is easier for us_i to cheer ourselves_i up after the day is over.’

c. “[...] a gente_i acaba nos_i iludindo”
 A GENTE end-up.3SG REFL.1PL deceive.GER
 ‘We_i end up deceiving ourselves_i.’

d. “[A]qui a gente_i consegue nos_i conectar [...]”
 here A GENTE can.3SG REFL.1PL connect.INF
 ‘Here we can connect [...].’

¹⁰ It is also not the case that 1PL agreement on the verb blocks the underspecified reflexive form *se*:

(i) “[A] gente_i se_i conhecemos através de amigos [...]”
 A GENTE REFL met.1PL through of friends
 ‘We_i met each other_i through friends [...]’

There is independent evidence showing that the object clitic *nos* functions as a formality marker in BP. In an attitude study, Freitag et al. (2018) confirm that *nos* is perceived as more formal than *a gente*. Schwenter et al. (2022) also show that clitic forms like *nos* are viewed as more *educated* and *intelligent* than their non-clitic counterparts (e.g., *nós* and *a gente*). It is widely attested that BP is gradually losing its object clitics, which may indicate that *nos* is seen as *conservative* (Carvalho & Calindro 2018; Cyrino 2003). All these properties relate to the central notion of *formality*.

What is most interesting is that configurations where *a gente* locally binds *nos* are much less frequent than cases like (13), where *a gente* locally binds *se*.¹¹ Table 4 contains an overview of the attested local agreement targets (for VA, PA and AA) found in connection to *a gente* (minus the dialectal 3PL). Some of these usage preferences are strong to the point that they have been argued to reflect grammatical constraints. For instance, some prior work has claimed on the basis of informal judgments that local agreement between *a gente* and *nos* is ungrammatical (Costa & Pereira 2013; Menuzzi 2000; Reuland 2011).

Table 4. Summary of local agreement patterns with *a gente*

	Consistently Attested Targets	Most Frequent Target
Verbal Agreement	3SG 1PL (3PL arguably dialectal)	3SG
Predicative Agreement	SG PL (MASC or FEM depending on reference)	SG (MASC or FEM depending on reference)
Local Anaphoric Agreement	underspecified (<i>se</i>) 1PL(<i>nos</i>)	underspecified (<i>se</i>)

As discussed above, speakers do in fact perceive the least frequent targets in Table 4 as marked or less acceptable compared to the most frequent ones. However, the empirical studies we cited suggests that, while the distributions of each variant differ across dialects, the variation in Table 4 itself exists within dialects and individuals as well. So speakers (to varying extents) *produce* these structures regardless of how they are judged. Furthermore, even those who reject the least frequent targets find them more acceptable than low acceptability baselines, such as cases where *a gente* triggers 2SG agreement:

(16) a. * A gente corre.
 A GENTE run.2SG
 ‘We run.’

 b. * A gente_i te_i viu.
 A GENTE REFL.2SG saw.3SG
 ‘We_i saw each other_i.’

From our perspective, the fact that the less frequent targets in Table 4 are (i) consistently attested within dialects/corpora, and (ii) more acceptable than blatant anomalies

¹¹ As a rough indication of this, we note that a search for “a gente nos_p VERB” vs. “a gente se_p VERB” in the BP section of *Corpus do Português: Web/Dialects* yields around 30 unique results for the former and over a thousand for the latter. However, the frequency of “a gente se_p VERB” is overestimated due to annotation errors for the part of speech tag “_p” in connection to the reflexive *se*.

like (16) suggests that they are not ungrammatical per se. However, the reduced acceptability of such targets and their relative sparsity does require an explanation. Descriptively, what is common to all of them is the presence of an overtly marked PL ϕ -feature. In Section 5, we propose an account of this in terms of a formal theory of register.¹²

3. Previous accounts

In this section, we review two proposals concerning the agreement properties of *a gente*: Costa & Pereira (2013) and Taylor (2009). The former was originally devised with European Portuguese in mind, but it is based on observations about BP made by Menuzzi (2000). These accounts posit that 1PL and 3SG agreement with *a gente* are triggered by distinct underlying structures – each with a different set of privative features acting as a local controller for agreement. Both accounts syntactically decompose *a gente* and derive the agreement variation in structural terms. In this respect, they are similar to accounts of variable agreement with collective *committee*-type nouns in English (Den Dikken 2001; Sauerland 2004a,b). We argue that such theories, which assume a one-to-one mapping between structures and agreement exponents, are too restrictive when it comes to the behavior of *a gente*. Ultimately, what has to be abandoned is the idea that the agreement information of a DP is represented as a single unstructured set of privative features.

3.1. Costa & Pereira (2013)

Costa & Pereira (2013) argue that *a gente* is a pronoun that can have two underlying structures. The variant giving rise to standard 3SG agreement is (17a), where *a gente* is the D-head of a 3SG DP. The variant giving rise to 1PL agreement is the complex nominal apposition structure in (17b), where *a gente* is adjoined to a 1PL empty pronoun. The latter is inspired by the structure Den Dikken (2001) ascribes to plurinals.

(17) a. $[_D P \ a \ gente_{3SG} \ NP]$ b. $[_D P \ pro_{1PL} \ [_D P \ a \ gente_{3SG} \ NP]]$

Building on prior work (Costa et al. 2001; Costa & Pereira 2005; Menzzi 2000), Costa and Pereira assume a phase-based modular approach to agreement, where the operation Agree is strictly local and can only access morphosyntactic ϕ -features (Chomsky 2000, 2001, i.a.). Semantic agreement is framed as a kind of last resort that applies when conditions for syntactic Agree are not met. This approach fundamentally relies on the distinction between *local* and *non-local* agreement, leading to the following predictions:

(i) non-locally: *a gente* should control 1PL, consistent with its semantics;

¹² A reviewer suggests that the less frequent agreement patterns in Table 4 are ungrammatical but may be repaired by ‘hypercorrection’, and, thereby come to sound more acceptable for some speakers. The main problem with this proposal is that there is no established theory of repair that derives why only some ungrammatical structures come to be regarded as acceptable. Notions like ‘hypercorrection’ are often invoked descriptively, but they have not been given an implementation in formal theories of grammar; as such, they are not explanatory mechanisms. To turn this intuition into a precise theoretical proposal one would need a model that (a) specifies how speakers identify members of the complement of the weakly generated set of grammatical objects, (b) defines a repair metric that licenses the particular attested repairs by comparing ungrammatical sentences to grammatical ones (perhaps along the lines of Chomsky (1961)), and (c) predicts the systematic patterning of those repairs across some domains (1PL agreement with *a gente*) but not others (e.g. 2SG agreement with *a gente*). As far as we are aware, this work has not been done.

(ii) locally: *a gente* can control 3SG or 1PL, depending on how it is parsed, cf. (17).

The first of these predictions is correct if we say non-local agreement is always *anaphoric agreement* (AA) (Bresnan & Mchombo 1987). Menuzzi (2000) notes that non-locally, *a gente* only binds forms with exclusive 1PL semantics: *nos/nós* or *a gente*.

(18) Só *a gente_i* sabe [que ele {*viu *ela_i* / *viu *elas_i* / nos_i viu / viu *a gente_i*}].
only A GENTE know that he saw her saw them 1PL saw saw A GENTE

However, Costa and Pereira assume that, in the structures underlying predicative agreement (PA), represented in (19), the relationship between the agreeing predicate and the subject is also not a local one, because the intervening SC counts as a separate phase.¹³

(19) [TP Subject₁ T⁰ [VP V⁰ [SC *t₁* Predicate]]]

Hence, for Costa and Pereira, there are cases of non-local agreement that are not anaphoric. One difficulty with this view is the fact that, as we saw in Section 2.2, the most common target for PA with *a gente* in BP is SG – i.e. what Costa and Pereira would take to be an instance of *syntactic agreement*. If PA is in fact non-local, this would be a counterexample to their analysis.

Costa and Pereira's prediction regarding *a gente*'s behavior in local domains is also incompatible with data in Section 2. In order to account for the variation in local agreement with *a gente*, they posit two distinct underlying structures: (17a), for 3SG, and (17b), for 1PL. However, once one of these structures is in place, all locally agreeing forms should be consistent with it. Costa & Pereira (2013: 167) thus claim that *a gente* can only trigger AA with the 1PL reflexive *nos* when it simultaneously triggers 1PL agreement with the verb. When VA targets are 3SG, only *se* is said to be appropriate:¹⁴

¹³ Costa & Pereira (2005) offer two arguments for this view. First, they claim that SCs constitute “a complete domain of predication” and “are prosodically coherent”, in that multiple SCs can be coordinated, as in (i), which they take to be properties of phases. Second, they claim that SCs, unlike nominal objects, cannot undergo leftward scrambling, as (ii) illustrates. This follows if scrambling is phase-internal and cannot access parts of phases that have been previously constructed and shipped to the interfaces.

(i) Eu considero [a Maria inteligente], [a Olga elegante] e [o Pedro irritante].
I consider the Maria smart the Olga elegant and the Pedro irritating
I consider Maria pretty, Olga elegant and Pedro irritating.

(ii) a. O Pedro fala francês sempre.
the Pedro speaks French always

b. * O Pedro está cansado sempre.
the Pedro is tired always

We are not convinced by either of these arguments. The first points to features that are necessary properties of phases, but not sufficient ones. TPs, for instance, are also complete domains of predication, but not phases. The ability to be coordinated is a general property of phrasal constituents (VPs, DPs, PPs, etc.) and also not a diagnostic for phasehood. About the second argument, the judgments seem not to be so robust and are potentially influenced by prosodic factors of the kind described in Menuzzi & Mioto (2006).

¹⁴ Costa and Pereira's judgments are given for European Portuguese, but Carvalho et al. (2018: 128-129) contend that they also hold for BP. The judgment that the AA pattern (20b) is unacceptable is also widely found in the literature on BP (Menuzzi 2000: 210, Reuland 2011: 131, Marcotulio et al. 2013: 127, i.a.).

(20) a. A gente_i nos_i vimos na TV.
 A GENTE REFL.1PL saw.1PL on-the TV

b. *A gente_i nos_i viu na TV.
 A GENTE REFL.1PL saw.3SG on-the TV

c. A gente_i se_i viu na TV.
 A GENTE REFL saw.3SG on-the TV

d. *A gente_i se_i vimos na TV.
 A GENTE REFL saw.1PL on-the TV

However, examples like (20d) can be found in corpora, as (21) shows:

(21) “[...] a gente_i se_i conhecemos através de amigos [...]”
 A GENTE REFL met.1PL through of friends
 ‘We_i met each other_i through friends.’

Furthermore, as we saw in (15), examples analogous to (20b) are also consistently attested in corpora. (22) is interesting because *a gente* binds *se* clitic and *nos* in the same sentence, making it clear that the choice between them is not a matter of dialect variation:

(22) “[A] gente_i se_i gostava muito, sempre que a gente_i nos_i
 A GENTE REFL liked.3SG much, every-time that A GENTE REFL.1PL
 cumprimentava, eu sentia que ela gostava de mim.”
 greeted.3SG, I felt that she liked of me
 ‘We_i liked each other_i, every time we_i greeted each other_i, I felt she liked me.’

While the markedness of structures where *a gente* locally binds *nos* requires explanation, adopting an account that entails ungrammaticality fails to account for why structures like (15) and (22) are consistently attested and judged to be *more* acceptable than blatantly deviant structures (e.g. *a gente_i te_i viu*). In the *nos* clause in (22), *a gente* triggers local VA based on syntactic ϕ -features and local AA based on meaning, showing that local agreement is not always determined by a unique set of syntactic ϕ -features. This is a pattern attested for hybrid NPs cross-linguistically, even in English (Smith 2021):

(23) The government is embarrassing themselves.

We also want to argue for a stronger claim: namely, that AA with *a gente* is never based on its syntactic ϕ -features. This is not only the case for non-local AA (which Costa and Pereira correctly predict), as we saw in (18), but also for intra-clausal binding, including binding of inalienable possessives and complex anaphors (Branco & Marrafa 1999; Varaschin 2021a), as in (24). The appropriate way to express binding with *a gente* in these and other cases is either with *a gente* itself or with overtly 1PL forms, as in (25).

(24) a. *A gente_i ama as {mães dela_i / suas_i mães}.
 A GENTE love.3SG the mothers of.her POSS.3 mothers
 ‘We_i love our_i mothers’

b. *A gente_i viu {ela mesma_i / elas mesmas_i} na TV.
 A GENTE saw her same.F.SG of-them.F same.F.PL on-the TV
 ‘We_i saw ourselves_i on TV.’

(25) a. A gente_i ama as {mães d-a gente_i / nossas_i mães}.
 A GENTE love.3SG the mothers of-A GENTE our mothers
 b. A gente_i viu {nós mesmas_i / a gente mesma_i} na TV
 A GENTE saw us same.FEM.PL a gente same.F.SG on-the TV

In fact, the only evidence Costa & Pereira (2013) and others (Menuzzi 2000; Reuland 2011) give in favor of the view that local AA with *a gente* is sensitive to purely syntactic ϕ -features are cases like (20c), where *a gente* locally binds the reflexive *se*. However, as we saw in (14), it is inappropriate to assign 3SG ϕ -features to *se* in BP. Therefore, the assumption that *a gente* can control 3SG AA locally when targets for VA are 3SG is also false. The crucial generalization concerning which features are available for agreement is not based on *locality*, but on whether the type of agreement is *anaphoric* or *grammatical*. In Section 4, we develop a system to capture this restriction.

3.2. Taylor (2009)

Taylor (2009) argues that, contrary to appearances, *a gente* is not a pronoun, but a DP with a rich and somewhat flexible internal structure. This flexibility is what enables him to derive a broader range of agreement variation than what Costa & Pereira (2013) do.

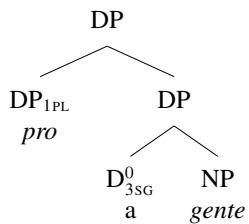
Taylor offers two arguments against the pronominal status of *a gente*. First, he claims that, unlike plural pronouns like *nós*, *a gente* cannot co-occur with numerals in structures like (26). Second, *a gente* cannot take NP complements in cases like (27).

(26) a. * A gente três foi à igreja.
 A GENTE three went to-the church
 b. Nós três fomos à igreja.
 we three went to-the church
 ‘We three went to church.’

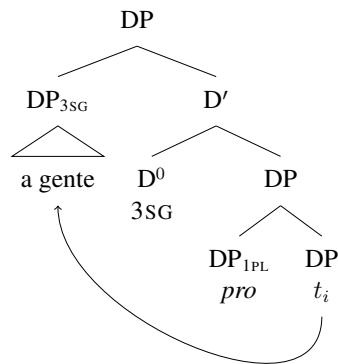
(27) a. * A gente brasileiros adora música.
 A GENTE Brazilians adore music
 b. Nós brasileiros adoramos música.
 we Brazilians adore music
 ‘We Brazilians adore music.’

Like Costa & Pereira (2013), Taylor (2009) also ascribes two structures for *a gente*, depending on whether it agrees with a 1PL or 3SG target. For 1PL uses, *a gente* has the appositional structure shown in (28a), similar to Costa and Pereira’s proposal. The key difference is that Taylor analyzes *a gente* as a full DP headed by a definite article, distinguishing it from true pronouns, which are D-heads (cf. Déchaine & Wiltschko 2002; Postal 1969, a.o.). This distinction explains why *a gente* cannot appear in numeral or nominal complement constructions like (26) and (27). The major innovation of Taylor’s approach is the idea that 3SG *a gente* is derived from the 1PL structure. The result is the imposter structure in (28b) (cf. Collins & Postal 2012), where the movement of the secondary source (i.e. *a gente*) to the Spec of the DP shell endows the higher D⁰ with 3SG features via Spec–Head agreement, which derives the 3SG patterns.

(28) a.



b.



Taylor's derivational account can derive local agreement mismatches like (29), which were wrongly ruled out under Costa and Pereira's proposal. The derivation works by assuming that the structure in (28a) is merged first, triggering 1PL agreement (with the anaphor or the predicative adjective), and subsequently moving to [Spec, TP]. Then, *a gente* is counter-cyclically adjoined to [Spec, DP], yielding the structure in (28b), which triggers 3SG agreement with the tensed verb. We illustrate this in (30) with a derivation for the mismatch between VA and AA in (29b) (see Taylor 2009: 23).

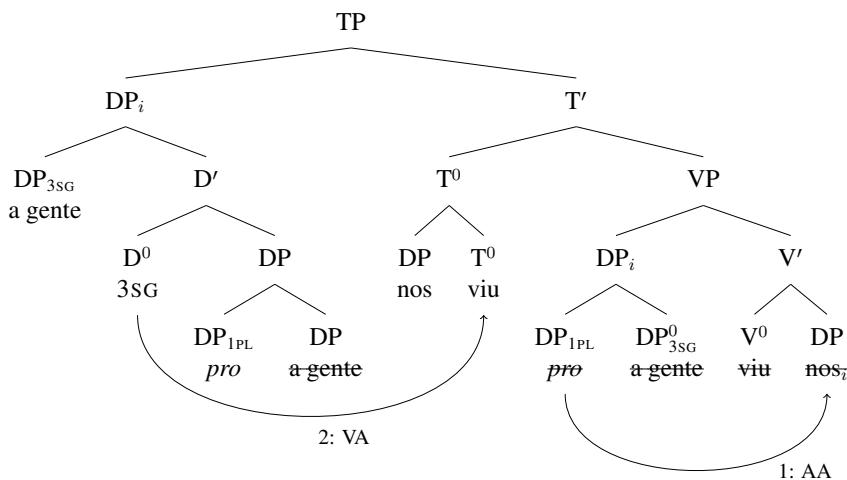
(29) a. A gente está cansados.
A GENTE is.3SG tired.M.PL

b. A gente_i nos_i viu.
A GENTE REFL.1PL saw.3SG

This account faces multiple challenges. First, neither of Taylor's arguments against the pronominal status of *a gente* arguments hold up under closer scrutiny (Costa & Pereira 2013; Marcotulio et al. 2013; Ramos & Maia 2015). Costa & Pereira (2013) show that the impossibility of (26a) and (27a) follows from two independent assumptions: (i) the requirement that determiners and modifiers of nominals agree with such nominals in their morphosyntactic ϕ -features; (ii) the fact that *a gente* is morphosyntactically specified as 3SG, cf. Section 4.¹⁵ Since numeral phrases and bare plurals like *brasileiros* are specified for PL as part of their morphosyntactic make-up, co-occurrence with the singular form *a gente* is ruled out. This has nothing to do with *a gente*'s status as a pronoun. Note that it is perfectly possible to use *a gente* followed by a singular bare noun, as the attested examples in (31) indicate.

¹⁵ Furthermore, not all pronouns in BP can take NP complements as in (27), as Carvalho (2018: 55) notes (see also Déchaine & Wiltschko 2002: 421). For instance, this is not possible for 1SG or 3PL forms: e.g. *eu brasileiro, *eles brasileiros.

(30)



(31) a. “A gente brasileir-o tem um preconceito danado contra a A GENTE Brazilian-M.SG have a prejudice damn against the internacionalizačo.”
 internationalization
 ‘We Brazilians have a lot of prejudice against internationalization.’

b. “Isso é uma injustiča contra a gente brasileir-o.”
 this is an injustice against A GENTE Brazilian-M.SG
 ‘This is an injustice against us Brazilians.’

Such cases are counterexamples to Taylor’s theory, since they should be blocked by whatever blocks multiple apposition in structures like **we linguists phonologists*. Also problematic are all of the pronominal properties of *a gente* mentioned in Section 2.1.

Second, though the derivational account of 3SG *a gente* can derive mismatches where *a gente* triggers PL agreement with a lower target and SG agreement with a higher one, the reverse order (i.e. where PL agreement is triggered high and SG is triggered low) is not derivable in Taylor’s system. However, we have seen that such mismatches are found (albeit marginally) in interactions between VA and PA, such as (11b) or (32).

(32) “[...] a gente tamos cansado [...]”
 A GENTE are.1PL tired.SG
 ‘We are tired.’

Third, Taylor’s account also cannot derive mismatches like (33). The first merge position of *a gente* in this case is subject of *triste*, where it should trigger SG PA. However, the adjective also takes the complex reflexive *n&osacute;s mesmas* as its complement, which targets *a gente*’s 1PL features. From this position, *a gente* moves to [Spec, TP], where it triggers 3SG agreement with T. There is simply no position where *a gente* could trigger, at the same time, SG PA with *triste* and 1PL AA with the reflexive.

(33) A gente_i t´a [t_i triste com n&osacute;s mesmas_i].
 A GENTE is.3SG sad.SG with 1PL same
 ‘We_i are sad about ourselves_i.’

Fourth, the fact that Taylor ascribes a normal non-pronominal DP structure to *a gente* fails to explain why the reading where *a gente* functions as a 1PL pronoun is lost if the head noun *gente* is modified, as in (34).

(34) A gente interessante cantou. . .
 A GENTE interesting.SG sang.3SG
 ‘The interesting people sang.’ (Not: ‘We sang.’)

Fifth, real imposter DPs do not allow bound indexical readings (Collins & Postal 2012: 195, Podobryaev 2014: 30). (35a) does not have a reading that entails that other people don’t think that they are brilliant, which (35b) does have: i.e. a reading where the second DP is abstracted over, yielding the property ‘ $\lambda x.x$ thinks *x* is brilliant’. However, both sentences in (36), including (36a), where 3SG *a gente* is bound, have this reading.

(35) a. Only we_i think that the present authors_{*i*} are brilliant.
 b. Only we_i think that we_i are brilliant.

(36) a. Só nós_{*i*} achamos que a gente_{*i*} é brilhante.
 only we think.1PL that A GENTE is.3SG brilliant
 b. Só nós_{*i*} achamos que nós_{*i*} somos brilhantes.
 only we think.1PL that we are.1PL brilliant

Lastly, the movement itself that is invoked in the derivation of 3SG *a gente* in (28b) arguably violates anti-locality constraints (Abels 2003; Boeckx 2008, i.a.).

3.3. General remarks

Costa & Pereira (2013) and Taylor (2009) disagree about the status of *a gente* as a pronoun and about the structure underlying examples where *a gente* controls 3SG agreement. However, they both agree that 1PL *a gente* has an appositive structure, where it is adjoined to an empty *pro* that bears the relevant 1PL features.

There are two problems that this shared assumption leads to. First, as Costa & Pereira (2013: 178–179) acknowledge, they fail to explain why similar instances of apposition are not possible for the overt 1PL pronoun *nós*. There is no reason why (37) should be bad, given that an identical structure is grammatical with the null 1PL *pro*.¹⁶

¹⁶ Costa & Pereira (2013: 179) stipulate, on the basis of (ia), that apposition to overt pronouns is independently ruled out. However, such examples are only marginally bad and more acceptable ones can be constructed, as (ib). It is apposition to *pro* – i.e. precisely the kind of apposition invoked to handle cases where *a gente* triggers 1PL agreement – that can be independently shown to be unacceptable, as in (ii).

(i) a. ?Eles, esses, cantam.
 they these.M sing.3PL
 b. Eu, eu-zinho, canto.
 I I-DIM.M sing.1SG

(ii) *Eu_{*i*} disse que *pro_{*i*}*, um linguista, posso ajudar.
 I said that *pro* a linguist can help
 ‘I_{*i*} said that, I_{*i*}, a linguist, can help.’

(37) * Nós, a gente, cantamos.
we A GENTE sing.1PL

Second, the presence of a null *pro* in all structures where *a gente* triggers 1PL agreement does not appear to conform to the constraints on the distribution of null pronouns in BP, which are known to be much stronger than in European Portuguese (Barbosa 2019; Kato & Duarte 2014; Soares et al. 2020, i.a.). For instance, referential null subjects in root declarative clauses are marginal in BP. However, there seems to be no restriction on 1PL *a gente* in similar contexts. Similarly, BP does allow 1PL *pro* in imperatives, but there is no corresponding imperative *pro* for 3SG *a gente*; i.e. (38b) does not have the same intended reading as (38a).

(38) a. Vamos sair!
go.1.PL.PRS leave-INF
'Let's leave!'
b. * Vai sair!
go.3.SG.PRS leave-INF
'Let's leave!'

The two approaches examined in this section also agree on the received view on ϕ -features, which takes them to be assembled in unstructured sets of interpretable privative elements (Adger 2006; Adger & Svenonius 2011; Collins & Stabler 2016). This yields predictions that are, at the same time, too restrictive and not restrictive enough.

On the former side, Costa and Pereira and Taylor's theories are incapable of explaining cases where *a gente* controls 3PL agreement, such as (9) above. It appears that what is going on there is that PERSON features are coming from *a gente*'s morphosyntactic make-up, while NUMBER is coming from *a gente*'s semantic information.

In another sense, both theories are also not restrictive enough. Insofar as both ascribe unique 3SG features to variants of *a gente* associated with 3SG verbal targets and *se*, they fail to capture the fact that cases of local AA with unambiguous 3SG forms, such as (24), are not possible. That is, they miss the generalization that AA with *a gente* is only sensitive to its referential features, regardless of whether it is local or not.

The alternative we propose in the next section remedies both of these defects. It assumes a richer theory of features that allows recursion – i.e. a feature can take sets of features as their values, one of which is not directly associated with interpretive effect at the truth-conditional level, but, nonetheless, may carry social meanings that determine its distribution across different types of contexts (e.g. formal or informal). Verbs and predicates can target their controller's syntactic and semantic features alike. But AA is stated in such a way that it requires sharing of semantic features.

4. Analysis: The grammar of *a gente*

In Section 3, we showed that, the hypothesis that ascribes a pronominal status to *a gente* is essentially correct. Nevertheless, previous attempts to define its agreement properties fail to capture the full range of attested variation, being mostly too strict and ruling out structures attested in the data. In this and the next section, we argue for an approach

that predicts – within a single grammar and with a single underlying structure for *a gente* (contra Taylor 2009; Costa & Pereira 2013) – the distributional facts found in BP corpora.

In our analysis – following Kathol (1999); Wechsler & Zlatić (2003); Wurmbrand (2017); Smith (2021); a.o. – we propose two sets of ϕ -features. One set, which we call CONCORD features (ϕ^C), contains the grammatical features needed a.o. for determiner–noun agreement and for the realization of morphophonological properties of inflectional heads (cf. Adamson 2025; Kathol 1999; Smith 2021).¹⁷ The other set, which we call INDEX features (ϕ^I), is more semantically oriented and functions to restrict the range of possible referents of an expression for referential pronouns as well as to constrain AA. We propose that tokens of ϕ^I -features replace traditional referential indices, so that they can, at once, govern particular forms of agreement, and encode semantic relations such as coreference and trigger binding via predicate abstraction (Heim & Kratzer 1998).¹⁸

Usually, ϕ^C - and ϕ^I -values of expressions are token identical, cf. (39a). But in some cases, for instance, for hybrid nouns (cf. Corbett 2015; Smith 2021), values can differ, cf. (40a), leading to competing agreement patterns, cf. (40b) vs. (39b).

- (39) a. Peter_{{ ϕ^C :{3,SG}, ϕ^I :{3,SG}}}
- b. Peter {sleeps / *sleep}.
- (40) a. committee_{{ ϕ^C :{3,SG}, ϕ^I :{3,PL}}}
- b. The committee {has / have} made a decision.

The pronouns *nós* and *a gente* – used to refer to a plurality including the speaker – offer a similar picture as (39a) and (40a): while the ϕ^C - and ϕ^I -values of *nós* are identical (41a), the values of *a gente* differ (41b). As mentioned in Section 1, the pronominal form *a gente* ‘we’ developed from the hybrid definite NP *a gente* ‘the people’, a hybrid noun like (40a), whose features also differed regarding the number values, cf. (41c).

- (41) a. ‘we’: *nós*_{{ ϕ^C :{1,PL}, ϕ^I :{1,PL}}}
- b. ‘we’: *a gente*_{{ ϕ^C :{3,SG}, ϕ^I :{1,PL}}}
- c. ‘the people’: *a gente*_{{ ϕ^C :{3,SG}, ϕ^I :{3,PL}}}

When used as a pronoun, the ϕ^I -values of *a gente* are specified as 1PL since ϕ^I -features are semantically interpretable and restrict the range of possible referents for the expression via general constraints on assignment functions, which will include at least conditions such as those in (42) (Sudo 2012; Varaschin et al. 2025). (42a) requires assignment functions to map a first person ϕ^I to entities that have speaker parts (possibly s_c itself) and (42b) requires assignment functions to map a plural ϕ^I to pluralities. λ -abstraction, which gives rise to assignment-independent bound-variable readings, is only licensed when binder and binees have identical ϕ^I s.

(42) *Constraint on Assignment Functions*

A function g from the set of ϕ^I to D_e is an admissible assignment function for a complete utterance u in a context c iff for every ϕ^I :

¹⁷ In contrast to Wechsler (2021: 228), our data suggest that ϕ^C also bears PERSON values.

¹⁸ This approach takes inspiration from theories that posit complex indices in order to solve the puzzles connected to indexical binding, which is only licensed under AA (Podobryaev 2014; Sudo 2012, i.a.).

- a. $s_c \sqsubseteq g(\{\phi^I: \{1\}\})$
- b. $g(\{\phi^I: \{PL\}\}) \in D_e - \text{ATOM}(D_e)$

Let us now examine the individual instances of agreement to assess how the two sets of features account for the observed patterns. Beginning with agreement within the nominal domain, as illustrated in (43a) and (43b), we observe that determiner – in this case, the pronouns *a gente* or *nós* – must agree with the head noun with respect to their ϕ^C -values, i.e. SG in (43a) and PL in (43b). So NP agreement seems to be ϕ^C -based.¹⁹

(43) a. $\{Nós / *A\ gente\}$ brasileir-os jogamos futebol.
 we A GENTE Brazilian-M.PL play football
 b. $\{ *Nós / A\ gente\}$ brasileir-o jogamos futebol.
 we A GENTE Brazilian-M.SG play football
 ‘We Brazilians play football.’

Regarding VA and PA, (44a) shows that subject and verb can agree either in CONCORD or INDEX (with the latter sounding more marked); and (44b) illustrates the same behaviour with respect to predicative adjective and subject agreement.

(44) a. *A gente* $\{\dot{e} / ?\text{somos}\}$ um time.
 A GENTE be.3SG be.1PL a team
 ‘We are a team.’
 b. *A gente* tá $\{\text{triste} / ?\text{tristes}\}$.
 A GENTE be.3SG sad.SG sad.PL

The contrast between (43) and (44) indicates that the fundamental agreement distinction cannot be attributed to locality, since the relation between determiner and noun as well as the relation between subject and verb (or *mutatis mutandis*, under standard assumptions, between subject and predicative adjective) is local. Consequently, it cannot be maintained that local agreement necessarily corresponds to ϕ^C -agreement.

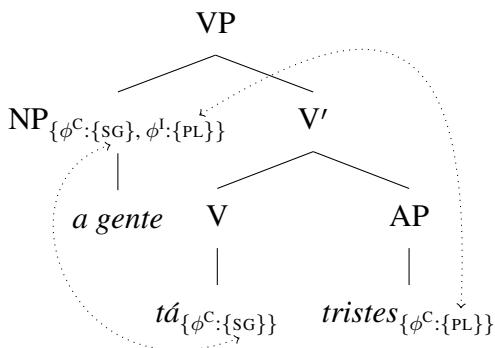
As previously noted, ϕ^C -features are employed in the realization of morphophonological properties; e.g. to determine the inflected form of a verb or predicative adjective. The rule in (45a), for instance, specifies that a head H^0 bearing the root specification α (e.g. *jog-* ‘play’) and the ϕ^C -value β (e.g. 1PL) is realized as the value of a paradigm function Φ that, when applied to a particular $\{\sqrt{ }, \phi^C\}$ bundle, determines the corresponding phonological form (e.g. *jogamos* ‘(we) play’) (cf. Kathol 1999; Stump 2001). The ϕ^C -value of the verb, in turn, agrees either with the ϕ^C - or the ϕ^I -values of the elements it enters into an agreement relation with. Thus, as illustrated in (45b), the verb can agree with *nós* (whose ϕ^C - and ϕ^I -values are 1PL) or with *a gente* (whose ϕ^I -value is 1PL).

(45) a. $H^0 \left\{ \begin{array}{l} \sqrt{ } : \alpha \\ \phi : \beta \end{array} \right\} \Rightarrow \Phi(\alpha, \beta)$
 b. *jog-amos* _{$\phi^C: \{1, \text{PL}\}$} selects $\text{NP}_{\phi^C: \{1, \text{PL}\}} \vee \text{NP}_{\phi^I: \{1, \text{PL}\}}$
 play-1PL

¹⁹ For independent reasons (e.g. the occurrence of bare singulars), we assume that BP belongs to the so-called NP languages (or *head-functor*-languages, cf. Bošković, 2007; Deng et al., 2025), in which the determiner – in this case the pronouns *a gente* or *nós* – selects the head noun.

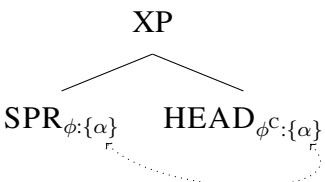
More intriguing are cases such as (44b), in which the ϕ -features of the verb and the predicative adjective do not match – for example, $tá_{\phi^C:\{3, SG\}} tristes_{\phi^C:\{PL\}}$ ‘is sad’. Such instances provide compelling evidence for a more articulated theory of features like the one we propose (cf. Lopes, 1999 for a similar approach to *a gente*), which distinguishes between CONCORD and INDEX-type ϕ -features. A theory assuming only one set of ϕ -features would be unable to account for this mismatch, as it would require the verb and the predicative adjective to share the same NUMBER value, i.e. both being either SG or PL, as discussed in Section 3. By contrast, the theory advocated here accounts for these data by allowing *a gente* to satisfy the constraints imposed by *tá tristes*²⁰ (cf. (46)): it is SG in ϕ^C or ϕ^I , and it is PL in ϕ^C or ϕ^I . $tamos_{\phi^C:\{1, PL\}} triste_{\phi^C:\{SG\}}$ is, thus, also licensed.

(46)



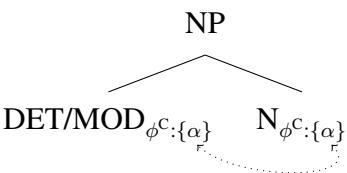
The cases involving VA and PA suggest a broader generalization about the syntactic configuration of agreement: in specifier-head configurations, a target bearing α ϕ^C -values selects a specifier whose values match either its ϕ^C - or ϕ^I -values, as illustrated in (47). We use ϕ as the underspecified form for ϕ^C and ϕ^I .

(47)



In contrast, within the nominal domain, the situation differs, as shown in (43). Within NPs,²¹ both determiner–noun as well as modifier–noun relations require agreement in their ϕ^C -values, as illustrated in (48). So the type of ϕ -feature that controls agreement in this case is not underspecified, but is determined to be ϕ^C .

(48)



²⁰ We assume a standard raising analysis for the combination of the verb with the predicative adjective, according to which the selectional features of the subject of the AP and those of the subject of the verb must be compatible. This can be enforced either by movement (from Spec, AP to Spec, VP) or structure-sharing.

²¹ We use ‘NP’ in a theory-neutral sense, irrespective of whether one adopts a DP- or NP-hypothesis. Our claim is intended to be more general and can be implemented within either theoretical framework.

In contrast to agreement in the nominal domain as well as in VA and PA, agreement involving anaphora and pronouns, i.e. AA, is not (necessarily) local. While anaphors are always in a local relation with their controllers (49a), pronouns are not (50) and can agree with their antecedents even across sentence boundaries. What is particularly noteworthy about AA is that *a gente* does not agree in ϕ^C (i.e. 3.SG.F) with its target, neither in local (cf. (49a) vs. (49b)) nor in non-local environments (cf. (50)).

(49) a. A gente_i {*se_i* / ?*nos_i*} viu na TV.
 A GENTE REFL 1PL saw.3SG on.the TV.
 INTENDED: ‘We_i saw ourselves_i on TV.’

b. *A gente_i viu *ela* *mesma_i* na TV.
 A GENTE saw.3SG 3.F.SG same.F.SG on.the TV.
 INTENDED: ‘We_i saw ourselves_i on TV.’

(50) a. A gente_i acha que a Ana gosta {de *nós_i* / **dela_i*}
 A GENTE think.3SG that the Ana likes.3SG of us of.her
 ‘We_i think that Ana likes us_i.’

b. A gente_i abraçou {*nossas_i* mães / **as mães dela_i*}
 A GENTE hugged.3SG our mothers the mothers of.her
 ‘We_i hugged our_i mothers.’

These data strongly suggest that, in the case of AA, only the more semantically oriented ϕ^I -agreement is required. This follows from the general properties of indices, namely their function in restricting the set of possible referents for an expression (in the case of free pronouns). That is to say, *a gente* and another expression – whether an anaphor or a pronoun – can be coreferential iff they refer to the same set of possible referents, a set restricted by their ϕ^I -values and not by their ϕ^C -values. Coindexing is also responsible for triggering binding via predicate abstraction (Heim & Kratzer 1998), in which case it is constrained by the principles of Binding Theory, which also independently determine whether an anaphor or a pronoun must be realized.²²

²² A reviewer raises the question of whether this approach predicts the behavior of epicene nouns. We are actually not making any claims about whether anaphoric GENDER agreement is also based on INDEX, or whether it follows from something else. This is a fascinating question we cannot address here for reasons of space. The answer partially hinges on the interpretation of examples like (i) when the victim is male.

(i) a. Só a vítima_i achou que ela_i ia morrer.
 only the.FEM witness.FEM thought that she would die
 Only the victim_i thought she_i was going to die.

b. Só a vítima_i achou que ele_i ia morrer.
 only the.FEM witness.FEM thought that he would die
 Only the victim_i thought he_i was going to die.

The judgments here are actually not clear to us. If only (ia) is acceptable under a bound-variable reading (the reading where other people did not think they would die), then we can assume that indices bear GENDER features along with PERSON and NUMBER. If both examples have this reading, one could assume, following Sudo (2012), that GENDER is actually not part of ϕ^I , but encoded as a presupposition on a separate head, along the lines of Heim & Kratzer (1998). An alternative would be to argue that the indices of epicene nouns like *vítima* (‘victim’)

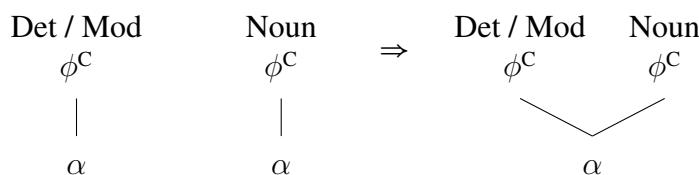
We have also observed that *a gente* agrees as well with the reflexive *se*, cf. (51a). However, as discussed in connection with (14), *se* can also co-occur with other pronouns, such as *tu* ‘you’, or, as illustrated in (14b) (repeated here as (51b)) with *nós* ‘we’. We therefore assume that the ϕ^I -values of *se* are not inherently specified as 3SG, but are instead underspecified. According to Principle A of the Binding Theory, since *a gente* and *se* must be coindexed, the ϕ^I -values of *se* are specified as 1PL.

(51) a. “A gente_i *se_i* gostava muito, sempre que a gente_i *nos_i*
 A GENTE REFL liked.3SG much, every-time that A GENTE REFL.1PL
 cumprimentava, eu sentia que ela gostava de mim.”
 greeted.3SG, I felt that she liked of me
 ‘We_i liked each other_i a lot, every time we_i greeted each other_i, I felt that she
 liked me.’

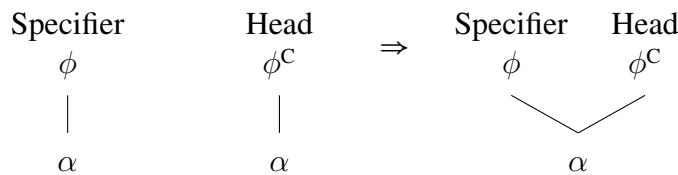
b. “Toda vez q[ue] nós_i *se_i* encontramos, nós oramos juntos.”
 all time that we REFL meet, we pray together
 ‘Every time that we_i meet each other_i, we pray together.’

Summarizing, the data in BP give rise to three agreement patterns that can be accounted for by assuming two different sets of ϕ -features: ϕ^C and ϕ^I . First, within the nominal domain, in the combination of determiner or modifier with noun, these elements must agree in their ϕ^C -values, as schematized in (52a).²³ Second, as the data for VA and PA agreement suggest, in specifier-head configurations, both elements agree w.r.t. their ϕ -values, i.e. with the CONCORD or INDEX specifications of the controller, cf. (52b). We represent this by structure sharing an underspecified ϕ , which can be resolved to either ϕ^C or ϕ^I . Third, since only ϕ^I plays a role in the semantic interpretation of pronouns, anaphors and pronouns agree with their controllers only in their ϕ^I -values, as (52c) shows.

(52) a. *Agreement Constraint within NPs*

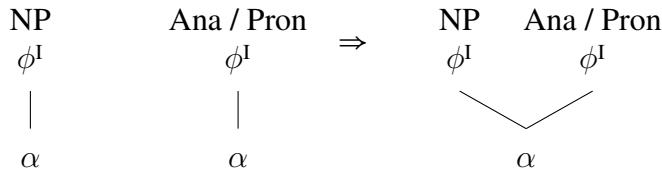


b. *Agreement Constraint for VA and PA*



are underspecified for GENDER, so the gender of the bound form actually resolves to the gender on the INDEX of its antecedent under AA. We leave this topic for a future study.

²³ Assuming a Head-Functor approach (cf. Van Eynde 2006) for NP languages – as proposed in Deng et al. (2025) – this constraint could be generalized since it would be stated in terms of a functor-head relation, similar to (52b) for specifier-head relations.

c. *Agreement Constraint for AA*

In this section, we have demonstrated that the fundamental theoretical distinction between the different agreement patterns cannot be framed solely in terms of local vs. non-local agreement. The BP data reveals a tension between two prevailing assumptions: first, that syntactic ϕ -features are active only in local domains; and second, that genuine semantic binding (as opposed to accidental coreference) is mediated by matching of ϕ -features (Bassi 2021; Kratzer 2009). Our analysis shows that these assumptions are only tenable if two sets of features are proposed: ϕ^C and ϕ^I . The advantages of this type of analysis are clear. First, unlike previous analyses (cf. Costa & Pereira 2005; Costa & Pereira 2013; Menuzzi 2000; Reuland 2011; Taylor 2009), it accounts for the different patterns attested in BP corpora without requiring the dismissal of such data as ungrammatical, cf. (15). Second, it can be extended to further cases of “quirky” agreement, such as hybrid agreement in BP (e.g. *a galera saíram* ‘the people left.3PL’) as well as in other languages (Bruening 2020; Corbett 2015; Den Dikken 2001; Smith 2021). Finally, insofar as NP-internal agreement is always ϕ^C -based and non-local AA is always ϕ^I -based, our analysis also captures the empirical patterns that underlie the Agreement Hierarchy, which posits that targets that are structurally closest to N should favor syntactic agreement, and those farthest away can favor semantic agreement (cf. Corbett 2006: 207).

5. Analysis: A compositional theory of social meanings

The dual-feature account proposed in Section 4 can license the full range of agreement variants controlled by *a gente* attested in the data examined in Section 2. As it stands, the account is somewhat unconstrained, because it predicts that all attested structures are equally grammatical, despite the fact that speakers have different judgments about their relative acceptability and markedness (cf. Table 4 above). So, for instance, structures like those in (53) are perceived as marginal to varying degrees.²⁴

(53) a. ? *A gente* corremos.
 A GENTE run.1PL
 ‘We run.’

b. ? *A gente* ’tá tristes.
 A GENTE be.3SG sad.PL
 ‘We are sad.’

²⁴ As we mentioned in Section 2, in spite of this, speakers have consistent judgments that all of the marked variants in (53a,c) are nonetheless more acceptable than the ones where *a gente* is associated with 2SG agreement targets (e.g. *a gente corres*, ‘we run.2SG’). We propose that cases like (53) and *a gente corres* are unacceptable for different reasons: only in the latter case are the reasons grammar-internal. We can distinguish, thus, three main levels of acceptability: fully acceptable (i.e. an utterance is fine on grammatical and extra-grammatical grounds), marked (i.e. an utterance is grammatically fine but extra-grammatically deviant) and plainly unacceptable (i.e. an utterance is not assigned a structure by the grammar).

c. ? A gente_i nos_i viu.
 A GENTE REFL.1PL saw.3SG
 ‘We_i saw each other_i.’

In this section, we will explain this by imposing further constraints on combinations involving *a gente* and different agreement targets. However, since the less-than-fully-acceptable and infrequent variants in (53) are nonetheless still possible, the nature of these constraints will not be syntactic, but, rather, part of a compositional theory of social meanings based on work in formal pragmatics (Machicao y Priemer et al. 2025; McCready 2019; Potts 2007; Taniguchi 2019; Varaschin et al. 2024, i.a.).

The basic idea we propose to account for the usage preferences in Table 4 is that structures where *a gente* is associated with overt 1PL exponents (be they verbal, predicative, or anaphoric) convey conflicting social meanings, making it difficult for speakers to infer what register they belong to. Building on Potts (2007), McCready (2010) and others, we model social meanings as gradable expressives and define the social meaning of a phrase as the intersection of the social meanings of its daughters (for each type of social meaning present in both). We assume that individual points in social meaning scales (i.e. social meaning degrees) stand for equivalence classes of contexts where an expression can be felicitously used with respect to that social meaning. This new approach to the formal modeling of variation entails that structures where (informal) *a gente* combines with (formal) 1PL have a very narrow intersection in their formality social meaning scale, accounting for their restricted distribution across all registers.

5.1. Social Meanings as Conventional Implicatures

By social meaning (SM) we understand any kind of non-at-issue content that indexes some socially-relevant property of the coordinates that define utterance contexts (e.g. s_c , a_c , t_c). Typical instances of SMs are (i) the expressive component of slurs: e.g. the word *Kraut*, which denotes the set of Germans and expresses the speaker’s negative attitude towards the kind corresponding to this set (Gutzmann 2011; McCready 2010); and (ii) politeness-related information conveyed by honorific forms: e.g. the form *du* in German, which refers to the addressee and expresses social or psychological closeness between the speaker and the addressee (Kim & Sells 2007; Potts & Kawahara 2004).

It has been shown that such SMs have projective and inferential properties associated with Conventional Implicatures (CI) (McCready 2010, 2019; Potts 2007 i.a.), The following are some of the properties of SMs we want to capture:

- Independence: SMs add meaning in a dimension separate from the at-issue content. Thus, the basic content of an utterance remains constant regardless of its SMs, and SMs are not affected by truth-conditional operators like modals or negation.
- Indexicality: SMs predicate something about the utterance situation or its participants, always referring to an individual or context contiguous with the utterance.
- Repeatability: Using multiple items conveying the same type of SM does not result in redundancy, but in specification of the contextual condition signaled by the SM.
- Immediacy: SMs realize their intended function through the act of utterance itself, they function like performatives, directly operating on the context of utterance.

- Gradability: The applicability of an SM is not absolute; rather, SMs apply to individuals to varying degrees, which may shift as a dialogue unfolds.

Note that the expressive content of slurs and the politeness content of honorifics have all of these features. Since these SMs share essential properties with CIs, particularly the Independence property mentioned above, they require a multidimensional approach to semantic interpretation. In this approach, different types of meaning contributions are separated into different levels: on the one hand, the at-issue dimension, and on the other, a CI-dimension, which can be thought of as a parameter of the context. A prominent approach of this sort that has been applied to SM phenomena like slurs and honorifics is that of McCready (2010) who, following Potts (2005), assumes a type-theoretic distinction between the two meaning dimensions: at-issue types are written with a subscripted ‘ a ’, and CI-types are written with a subscripted ‘ c ’. The compositional rules guarantee no CI-type can serve as an argument to an at-issue type, accounting for the independency property of SMs. The CI composition is ‘isolated’ from the at-issue meaning by means of the \bullet operator. On this approach, slurs and honorifics would be assigned denotations like those in (54).²⁵

(54) a. $\llbracket \text{Kraut} \rrbracket = \lambda x. \text{german}(x) : \langle e^a, t^a \rangle \bullet \text{finds-bad}(s_c, \cap \text{german}) : t^c$
b. $\llbracket du \rrbracket = a_c : e^a \bullet \text{low-distance}(s_c, a_c) : t^c$

In order to model the gradable property of SMs, we can modify these entries by requiring SMs to take an additional DEGR(EE) argument, along the lines suggested in Potts & Kawahara (2004), Potts (2007) and McCready (2019). In our approach, this will be an interval from 0 to 1, with 0 signaling the lowest degree and 1 the highest. So, for instance, we can think of *Kraut* as conveying that the speaker has a negative/low attitude toward the kind corresponding to the property **german** and of *du* as conveying that the speaker and the addressee gave a lower-to-average degree of social distance to each other:

(55) a. $\llbracket \text{Kraut} \rrbracket = \lambda x. \text{german}(x) : \langle e^a, t^a \rangle \bullet \text{attitude}(s_c, \cap \text{german}, (0, .3]) : t^c$
b. $\llbracket du \rrbracket = a_c : e^a \bullet \text{distance}(s_c, a_c, (0, .5]) : t^c$

The fact that we have continuous values reflects how we can make comparative judgments about SMs (e.g. speaker s_1 is more negative towards Germans than s_2) and intuitively sense varying degrees of SMs (e.g. expression e_1 conveys more social distance than e_2).

Recently this type of formalism has been extended to the treatment of register and situationally-driven intra-individual variation (McCready 2019; Taniguchi 2019; Varaschin et al. 2024). What this work argues for is that register parameters like *formality*, which certain expressions are conventionally associated with, can be understood as gradable SMs, much like those we see in (54)–(55). We assume a particular implementation of the SM-based approach to register developed in more technical detail in Varaschin et al.

²⁵ The original \mathcal{L}_{CI} logic in Potts (2005) did not allow expressions to contribute simultaneously to the at-issue and CI dimensions. The logic was extended by McCready (2010) and Gutzmann (2011) to allow for such kinds of mixed content. The entries in (54) are simpler than the ones suggested by McCready’s (2010) particular treatment of mixed content, because they keep the CI and at-issue contribution separate by means of the \bullet operator, as opposed to using shunting product types. The simplification is harmless in this context because the rules for mixed application essentially convert a complex shunting type into the types of semantic objects we see as the denotations in (54).

(2024). This approach incorporates the basic insights of McCready's system illustrated in (54)–(55). However, as opposed to using Potts's (2005) bullet ('•') operator, (AT-ISSUE) CONTENT is separated from CI meaning by assigning each to different attributes in a structured representation of the information conveyed by linguistic expressions, using the resources of sorted feature logics (Asadpour et al. 2022; Bender 2007; Green 1994; Paolillo 2000; Pollard & Sag 1994). This representation takes the general form in (56).

$$(56) \quad \begin{bmatrix} \text{CONTENT} & \text{at-issue meaning} \\ \text{CONTEXT} & \begin{bmatrix} \text{CONTEXTUAL-INDICES} & \langle s_c, a_c, t_c, \dots \rangle \\ \text{CI} & \langle SM_1, SM_2, \dots \rangle \\ \text{PRESUPPOSITION} & \langle P_1, P_2, \dots \rangle \end{bmatrix} \end{bmatrix}$$

The defining features of linguistic contexts are represented under the CONTEXT attribute. This includes the classic Kaplanian contextual parameters (under CONTEXTUAL-INDICES), a list of CIs (where SMs appear), and possibly also a list of PRESUPPOSITIONS. The logical form of SM propositions is exactly as it is in the right-hand side of the bullet in (55). SMs always take a degree argument, in addition to at least one of the members of the CONTEXTUAL-INDICES list (which captures their indexical nature).

One advantage of the approach to SMs we propose over the more standard system based on Potts (2005) is that it allows us to define the SM of an utterance compositionally. In technical terms, we can dispense with the non-compositional rule of parsetree interpretation (Potts 2005: 68) and adopt a compositional projection principle that defines the SM of a mother node in a semantic derivation as the append of lists of distinct SMs under the daughter's CONTEXT|CI and intersects the degree values of SMs of the same type (see Varaschin et al. (2024) for more details). The independence property of CIs is guaranteed by stating separate principles of composition CONTENT and CONTEXT. This ensures that operators over at-issue content (e.g. negation, modals, interrogatives) only pick out their scopal arguments from elements under CONTENT.

SMs are assigned to expressions by means of use-conditional constraints (UCCs), with the form in (57). The antecedent of a UCC specifies independently licensed structures on which the consequent imposes a contextual appropriateness condition, which we model as the SMs that have to be satisfied in order for the utterance to be felicitous. The antecedents can be lexical items (i.e. bundles of linguistic features), individual linguistic features, as well as possibly larger chunks of syntactic structures.

$$(57) \quad \text{linguistic structure } S \Rightarrow \text{admissible context for } S$$

In what follows, we propose UCCs associated with *a gente* and forms of variable agreement as well as an explicit principle of SM composition for calculating the formality of an utterance on the basis of the formality specification of its parts. These ingredients will allow us to derive the usage preferences for *a gente* documented in Section 2.

5.2. Accounting for the usage preferences

As we saw in Section 2, all of the variants examined in this paper are positively or negatively correlated with the situational parameters that are (more or less directly) associated with *formal(ity)*. In particular, we saw that *a gente* is the most frequent form of 1PL, but is

also perceived as less formal than *nós* (Brustolin 2009; Freitag 2016; Freitag et al. 2018; Freitas & Carvalho 2020; Lima 2015; Vianna 2006). In contrast, forms of overt PL inflection – which, as we proposed in Section 4, come about as externalizations of the $\{\phi^C: pl\}$ specification on heads – are all perceived as conveying a high degree of formality, which explains their association with written genres and older and more educated speakers (Alencar 2013; Brito & Sedrins 2017; Brustolin 2009; Freitag 2016; Zilles 2005). The clitic form *nos* arguably has a similar SM (Carvalho & Calindro 2018; Cyrino 2003; Freitag et al. 2018; Schwenter et al. 2022).

Following the sociolinguistic work we reviewed in Section 2, we assign the UCCs in (58) to account for SMs to the forms under consideration. We assume more specific SMs (i.e. SMs with narrower intervals) override less specific ones.²⁶

(58) a. *a gente* \Rightarrow [CONTEXT|CI $\langle \mathbf{formal}(s_c, (0, .8]) \rangle$]
 b. *nos* $\{\phi^I: Ipl, \phi^C: Ipl\}$ \Rightarrow [CONTEXT|CI $\langle \mathbf{formal}(s_c, [.7, 1]) \rangle$]
 c. $H^0\{\phi^C: pl\}$ \Rightarrow [CONTEXT|CI $\langle \mathbf{formal}(s_c, [.6, 1]) \rangle$]

Note that the inference that the speaker is presenting as **formal** when they employ a forms like *nos* or a verbal head specified for $\{\phi^C: pl\}$ in (59) has all of the SM properties listed in Section 5.1. This content is *independent* of the main point of the utterance (e.g. it would not be affected by negation); it is *indexical* (i.e. it says something about the speaker); there is no sense of redundancy in *repeating* forms that signal a similar degree of formality (*nos* and $H^0\{\phi^C: pl\}$); the effect of signaling high formality is *immediate* (i.e. it comes about simply by uttering the relevant expressions); and, finally, it is also *gradable* (e.g. it is conceivable that *formality* could be expressed to an even higher degree).

(59) Querem que nos odiemos.
 want.3PL that REFL.1PL hate.1PL
 ‘They want us to hate each other.’

As the UCCs in (58) already make clear, the SMs of *a gente*, on the one hand, and *nos* and $H^0\{\phi^C: pl\}$ on the other are somewhat at odds with each other because they pick out opposite ends of the scale corresponding to the **formal** SM. This raises the question of how these elements interact when they are combined in a single utterance. If we look again at Table 4, we will see that the less frequent and more marked patterns of agreement with *a gente* are precisely those where *a gente* combines with elements bearing $\{\phi^C: pl\}$ features (either *nos* or PL-inflected verbal and predicative heads).

The answer to this question is given by a general principle of SM composition, which Varaschin et al. (2024) call the Local CI Projection Principle (see also Asadpour et al. 2022; Machicao y Priemer et al. 2025; Paolillo 2000). This principle has different consequences depending on whether the SMs being combined represent the same situational

²⁶ We intend here merely to illustrate a general pattern of analysis that accounts for the markedness of certain variants in pragmatic terms, as opposed to offer a fully detailed account where all of the SMs are exhaustively investigated. So, as is usual in the literature on expressives and SMs, the particular intervals assigned as the degree arguments signal the overall direction of the observed effects. For the sake of simplicity, we also assume that all other linguistic structures under consideration (e.g. $H^0\{\phi^C: sg\}$, the other lexemes and constructions, etc.) are register neutral – i.e. they are specified as $\mathbf{formal}(s_c, (0, 1))$. A more realistic model would have to independently assess the SM contributions of these other variants.

parameter (e.g. both specify different degrees of formality) or different parameters (e.g. one specifies a degree of formality and another specifies the speaker's attitude towards a referent). Since we are dealing here solely with a single register-related SM type (i.e. **formal**) we adopt a simpler formulation applied specifically to the formality parameter.

(60) *Local Formality Composition Principle*

The DEGREE value for the **formal** SM of a phrase is the intersection of the DEGREE values of the **formal** SMs of its immediate daughters.

The intersection of DEGREE values δ_1, δ_2 is defined as the set of points that are in both δ_1 and δ_2 . A consequence of this principle is that the information content of CI values grows monotonically as more elements are combined. The **formal** value of a phrase is either maintained or further restricted, imposing tighter constraints on the contexts where the mother phrase can be appropriately used. This reflects the Repeatability property of SMs discussed above: repetition of SMs leads not to redundancy but to reaffirmation or specification, unlike the repetition of at-issue meanings. The principle in (60) also imposes a consistency requirement on SMs: the degree of formality of different elements must intersect, otherwise no CI value is defined for the mother. This is arguably what happens in instances of honorific clashes in languages like Punjabi or Thai, which have more grammaticalized politeness marking (Kaur & Yamada 2022; McCready 2019).

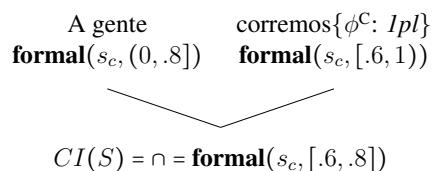
Register felicity is calculated by comparing the CI value of a sentence S that is the output of the Local Formality Composition Principle in (60) – $CI(S)$ – with the prior global context in which S is uttered (CX_S). This prior global context CX_S is, like $CI(S)$, also a list of SMs, but, for our purposes, we assume it to be a singleton list containing only a specification for the formality level of the prior discourse.

(61) *Felicity constraint* (cf. McCready 2019: 31)

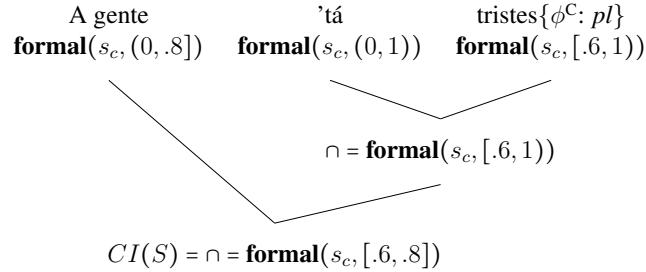
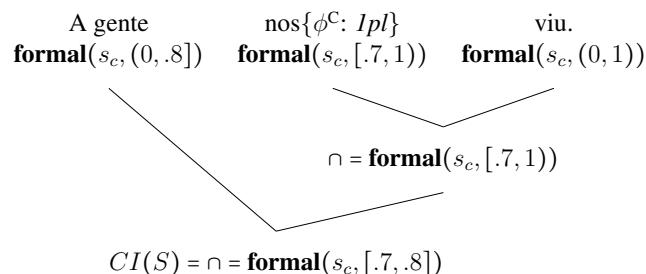
The DEGREE values of **formal** in $CI(S)$ and CX_S have to intersect.

All things being equal, the more the degree intervals of **formal** in $CI(S)$ and the one in CX_S match (i.e. the more degrees they share relative to their total size), the more register appropriate S is with respect to CX_S . With this in mind, we can understand why the agreement variants with *a gente* in (53) are perceived as marked, despite being licensed as grammatical by the dual-feature theory we proposed in Section 4. In Figures 2-4 we illustrate the effects of (60) for all of the marked agreement targets in (53).

Figure 2. Local Formality Composition with VA



As a result of the narrow intersection between the **formal** degrees between, on the one hand, *a gente* and, on the other, $\{\phi^C: pl\}$ heads and *nos*, the CI value of each full S will be such that the DEGREE value for the SM **formal** will be a very narrow interval. This means that these utterances are appropriate in a very narrow set of contexts – those

Figure 3. Local Formality Composition with PA**Figure 4.** Local Formality Composition with AA

in which there is a high, but not too high degree of formality. Since contexts with these precise characteristics are not particularly common, utterances like those in (53) will also be infrequent and will, thus, be dispreferred in more neutral situations, which are the kinds of situations that tend to be reflected in naive linguistic judgments.²⁷

There is one remaining puzzle – originally noted by Menuzzi (2000) – that we would like to address. This is the fact that local AA mismatches involving *a gente* are much more marked than non-local AA mismatches, which are fully acceptable on all accounts:

(62) a. ?A gente_i nos_i viu.
 A GENTE REFL.1PL saw.3SG
 ‘We_i saw each other_i.’

²⁷ The markedness of the targets in (i) follows from similar consideration. In all of these cases, *a gente* is combined with a plural element that shares with which it only shares a narrow overlap in formality.

(i) a. A gente ‘tamos tristes.
 A GENTE be.1PL sad.PL
 ‘We are sad.’

b. A gente ‘tamos triste.
 A GENTE be.1PL sad.SG
 ‘We are sad.’

c. A gente_i nos_i vimos.
 a gente REFL.1PL saw.1PL
 ‘We_i saw each other_i.’

b. A gente_i sabe que ele nos_i viu.
 A GENTE know.3SG that he REFL.1PL saw.3SG
 ‘We_i know that he saw us_i.’

The theory we formulated so far does not obviously predict this. Assuming all forms other than *nos* and *a gente* in (62) to be register neutral, the **formal** degree in the resulting CI(S) in (62b) would be the same narrow interval as the one we get in (62a).

We suggest an explanation of this inspired by the dynamic approach to register update in McCready (2019: 33). First, we postulate that the Local Formality Composition Principle in (60) is computed separately for each minimal S (i.e. CIs are computed at the CP level, excluding embedded CPs). Second, we propose that, if felicitous, the CI of each minimal S, $CI(S)$, automatically updates the global context CX_S . Third, we propose that the way contextual update works is different from how local CI composition works: update gives us the *average* of the DEGREE values for SMs like **formal** (not the intersection). It is, thus, non-monotonic. The relevant principle is stated in (63).

(63) *Global Contextual Update for Formality* (cf. McCready 2019: 33)
 If $CI(S)$ is felicitous, $CI(S)$ updates CX_S as CX'_S , where the DEGREE value of *formal* in CX'_S is the average between the DEGREE values of *formal* in $CI(S)$ and CX_S .

The average degree (μ_d) value for two intervals $[i_1, j_1]$ and $[i_2, j_2]$ is obtained by adding the maximum and minimum degree values of each and dividing them by two:

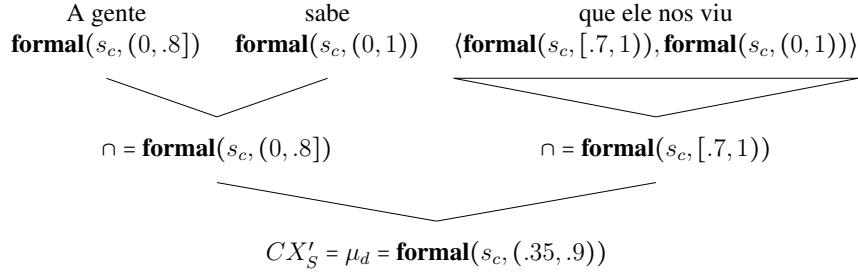
$$\mu_d([i_1, j_1], [i_2, j_2]) = \left[\frac{\min([i_1, j_1]) + \min([i_2, j_2])}{2}, \frac{\max([i_1, j_1]) + \max([i_2, j_2])}{2} \right]$$

What (63) says is that, if the output of Local Formality Composition for a sentence S is felicitous, it updates the global context, dynamically pushing it in the direction of the formality contribution of S . As McCready (2019) notes, this approach allows the degree of formality of a discourse to gradually shift as dialogue progresses.

Crucially, since Global Contextual Update gives us the average of the *formality* degrees for each minimal S, when *a gente* and *nos* are in different sentences, the resulting degree argument for the **formal** SM will not be the same narrow interval in Figure 4 (which is obtained by intersection), but the much broader interval in Figure 5. Since SM degrees represent equivalence classes of contexts, this means that structures with non-local AA between *a gente* and *nos* will be felicitous in a much broader range of situations than those where there is local AA between *a gente* and *nos*. We assume that this is why non-local AA is perceived as less marked/more acceptable than local AA variants.

6. Conclusion

This paper has offered an analysis of the agreement behavior of the first-person plural pronoun *a gente* in Brazilian Portuguese, a form whose grammatical properties have posed challenges for existing theories. We argued that prior accounts, which posit multiple syntactic structures corresponding to different agreement patterns triggered by *a gente*, make

Figure 5. Global Contextual Update with AA

incorrect predictions about the range of attested data, particularly in cases of agreement mismatches involving verbal, predicative, and anaphoric targets.

Instead, we proposed that *a gente* is a true pronoun with a hybrid ϕ -feature specification: semantically interpretable INDEX features (ϕ^I) valued as 1PL, and morphophonological CONCORD (ϕ^C) features valued as 3SG. This distinction accounts not only for the range of possible agreement targets but also for the grammatical asymmetries between agreement types – particularly the fact that anaphoric agreement tracks INDEX features exclusively. In our system, this follows from the fact that ϕ^I features take over the role referential indices in the Binding Theory. We further argued that the reduced frequency and low acceptability of agreement combinations involving overt 1PL morphology (e.g., *a gente* + *somos* or *nos*) are best understood not as grammatical violations, but as outcomes of social meaning composition. These combinations instantiate opposing register values, which we modeled as conventional implicatures that constrain utterance contexts.

Beyond the empirical domain of Brazilian Portuguese, our findings carry broader implications for the theory of ϕ -features. The most immediate implication is that they provide strong support for an articulated feature architecture in which ϕ -features are distributed across multiple dimensions (INDEX and CONCORD), with different syntactic and interpretive roles. This layered representation allows agreement to be sensitive to the demands of different targets, without requiring ad hoc or excessively restrictive structural stipulations. It requires, however, abandoning the simple architecture where ϕ -features are modeled as unstructured sets of interpretable privative elements (Adger 2006; Adger & Svenonius 2011; Collins & Stabler 2016). By contrast, the dual-feature system we adopt assumes that feature structures can be recursive; i.e. a single set of ϕ -features can itself include other sets of features representing its ϕ^I and ϕ^C values. While more complex, this architecture has also been independently argued to be necessary to capture hybrid agreement phenomena in different languages (Adamson 2025; Kathol 1999; Smith 2021; Wechsler & Zlatić 2003; Wurmbrand 2017, i.a.).

Perhaps less obviously, our analysis also supports a reinterpretation of the Principle of Radical Interpretability, which is sometimes assumed in discussions about Agree:

(64) *Principle of Radical Interpretability* (Brody 1997; Pesetsky & Torrego 2007)
 Each feature must receive a semantic interpretation in some syntactic location.

We showed that some features (INDEX features) constrain semantic interpretation at the level of at-issue content – these are the features that are typically taken to be ‘interpretable’. CONCORD features, in turn, primarily determine morphophonological realiza-

tion – they are, therefore, good candidates for being ‘uninterpretable’ in every syntactic location – especially when they do not coincide with a controller’s ϕ^I features.

However, we have also showed that even these typically uninterpretable features (ϕ -features on tensed verbs, predicative heads, etc.) can be use-conditionally interpretable – i.e. they can be used to convey social meanings about the speaker, the addressee, and other context coordinates. In this sense, ϕ^C features are interpretable not in terms of at-issue propositional content, but at the level of conventional implicature (Potts 2005).²⁸ This is a consequence of the fact that any aspect of linguistic structure – including aspects that do not have a ‘meaning’ in the at-issue sense – can be socially monitored in one way or another (Bender 2007; Meyerhoff & Walker 2013; Robinson 2022, i.a.).

This observation weakens the motivation for theories that require uninterpretable features to be deleted before transfer to the interfaces (cf. Chomsky 2000, 2001). If so-called uninterpretable features can be interpretable at the CI level, the dichotomy between interpretable and uninterpretable features becomes less stable. Our results suggest that instead of classifying features by whether they contribute to at-issue interpretation, we could classify them by the dimension of meaning they access: at-issue meaning, presupposition, use-conditions, or some other interface. In this way, Interpretability-type principles can be reformulated not as constraints against the presence of uninterpretable features, but as a condition on the grammatical architecture: every feature must be potentially interpretable somewhere – along some dimension of meaning. Perhaps all so-called ‘uninterpretable’ features can be interpretable only at non-at-issue dimensions of meaning like CI.²⁹

Finally, we would like to note that our analysis also raises broader methodological questions for grammatical theory. In particular, it provides another rationale for the view that acceptability judgments cannot be immediately interpreted as diagnostics for grammatical well-formedness (Culicover et al. 2022). An expression that is grammatically well-formed but specified with social meaning that is not fully felicitous in usual contexts (e.g. the combination of *a gente* with overt 1PL agreement, in our analysis) can often be indistinguishable from an expression that is ungrammatical in the narrow sense. That is, the psychological response from naive speakers may be the same in both cases. Such cases demonstrate that unacceptability can arise from problems at the level of register and social meaning, rather than from violations of core grammatical principles.

Acknowledgements

We would like to thank the audiences at Oberseminar Syntax and Semantics (Goethe-Universität Frankfurt am Main), ESSLLI 2023 (University of Ljubljana), the Berlin-BrnoVienna Workshop 2023 (Masaryk University), and Going Romance 2024 (Universidade do Minho), as well as our two anonymous reviewers. We profited from invaluable conversations with Luke Adamson, David Adger, Sigi Bahr, Pilar Barbosa, Carla Bombi, Aron Hirsch, Elin McCready, Andrés Saab, Manfred Sailer, Viola Schmitt, Frank Richter,

²⁸ This proposal is compatible with approaches that propose that register-related SMs are attached only to PF structures, such as Saab (2021).

²⁹ This could also be taken to support symmetric and feature-sharing approaches to agreement (e.g. Ackema & Neeleman 2013; Haug & Nikitina 2016; Kathol 1999), insofar as apparently ‘uninterpretable’ ϕ^C features may be shared across syntactic constituents without needing to be eliminated under valuation.

and Gert Webelhuth. All remaining errors are ours. This research was funded by the Deutsche Forschungsgemeinschaft (DFG) – SFB 1412, 416591334, Project A04.

References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. Storrs, CT: University of Connecticut Dissertation.

Ackema, Peter, & Ad Neeleman. 2013. Subset controllers in agreement relations. *Morphology* 23. 291–323. <https://doi.org/10.1007/s11525-013-9218-4>.

Adamson, Luke James. 2025. Split coordination with adjectives in Italian: An approach with multidominance and semantic agreement. *Natural Language & Linguistic Theory* 43(1). 3–62. <https://doi.org/10.1007/s11049-024-09617-5>.

Adger, David. 2006. Remarks on Minimalist feature theory and Move. *Journal of Linguistics* 42(3). 663–673. <https://doi.org/10.1017/S002226706004221>.

Adger, David, & Peter Svenonius. 2011. Features in Minimalist syntax. In Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, vol. 1, 27–51. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199549368.013.0002>.

Alencar, Leonel Figueiredo de. 2013. Modelação computacional de padrões variáveis de concordância em português. *Revista de Estudos da Linguagem* 21(1). 43–110. <https://doi.org/10.17851/2237-2083.21.1.43-110>.

Asadpour, Hiwa, Shene Hassan, & Manfred Sailer. 2022. Non-*wh* relatives in English and Kurdish: Constraints on grammar and use. In Stefan Müller, & Elodie Winckel (eds), *Proceedings of the 29th International Conference on Head-Driven Phrase Structure Grammar, Nagoya University & Institute for Japanese Language and Linguistics*, 6–26. Frankfurt am Main: University Library. <https://doi.org/10.21248/hpsg.2022.1>.

Barbosa, Pilar P. 2019. *pro as a minimal nP: Toward a unified approach to pro-drop*. *Linguistic Inquiry* 50(3). 487–526. https://doi.org/10.1162/ling_a_00312.

Bassi, Itai. 2021. *Fake feature and valuation from context*. Cambridge, MA: MIT Dissertation.

Bender, Emily M. 2007. Socially meaningful syntactic variation in Sign-Based Grammar. *English Language & Linguistics* 11(2). 347–381. <https://doi.org/10.1017/S1360674307002286>.

Boeckx, Cedric. 2008. *Understanding Minimalist syntax: Lessons from locality in long-distance dependencies*. Oxford: Blackwell.

Borges, Paulo R. S. 2004. *A gramaticalização de A gente no português brasileiro. Análise histórico-social-linguística da fala das comunidades gaúchas de Jaguarão e Pelotas*. Porto Alegre: Universidade Federal do Rio Grande do Sul Dissertation.

Bošković, Željko. 2007. What will you have, DP or NP? In Emily Elfner, & Martin Walkow (eds), *NELS 37: Proceedings of the Thirty-Seventh Annual Meeting of the North East Linguistic Society*, vol. 1, 101–114.

Branco, António, & Palmira Marrafa. 1999. Long-distance reflexives and the binding square of opposition. In Gert Webelhuth, Jean-Pierre Koenig, & Andreas Kathol (eds), *Lexical and constructional aspects of linguistic explanation*, 163–177. Stanford, CA: CSLI Publications.

Bresnan, Joan, & Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language*. 741–782. <https://doi.org/10.2307/415717>.

Brito, Dorothy Bezerra Silva de. 2008. Concordância e a relação antecedente/reflexivo no português brasileiro. *Revista do GELNE* 10(2). 42–45. <https://periodicos.ufrn.br/gelne/article/view/12157> (11 December, 2025).

Brito, Dorothy Bezerra Silva de, & Adeilson Pinheiro Sedrins. 2017. Sistemas de concordância e de correferência no paradigma pronominal de primeira pessoa do plural no português brasileiro. *Atas do V Simpósio Mundial de Estudos em Língua Portuguesa*. 13–28. <https://doi.org/10.1285/i9788883051272p13>.

Brody, Michael. 1997. Perfect chains. In Liliane Haegeman (ed.), *Elements of grammar: Handbook in Generative Syntax*, 139–167. Dordrecht: Springer. https://doi.org/10.1007/978-94-011-5420-8_3.

Bruening, Benjamin. 2020. The head of the nominal is N, not D: N-to-D movement, hybrid agreement, and conventionalized expressions. *Glossa: A Journal of General Linguistics* 5(1). 15. 1–19. <https://doi.org/10.5334/gjgl.1031>.

Brustolin, Ana Kelly Borba da Silva. 2009. *Itinerário do uso e variação de nós e a gente em textos escritos e orais de alunos do ensino fundamental da rede pública de Florianópolis*. Florianópolis: Universidade Federal de Santa Catarina MA thesis.

Cardinaletti, Anna, & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Volume 5/Part 1: Clitics in the Languages of Europe*, 145–234. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110804010.145>.

Carvalho, Danniell da Silva. 2018. *The internal structure of personal pronouns*. Newcastle: Cambridge Scholars Publishing.

Carvalho, Danniel da Silva, & Dorothy Bezerra Silva de Brito. 2017. Impostores, cor-referência e concordância em português brasileiro. *Revista Letras* 1(96). 54–73. <https://doi.org/10.5380/rel.v96i0.50421>.

Carvalho, Danniel da Silva, Dorothy Bezerra Silva de Brito, & Adeilson Pinheiro Sedrins. 2018. Referência de primeira pessoa e anáfora em português brasileiro. *ReVEL* 16(30). 127–145.

Carvalho, Janayna Maria da Rocha. 2021. Características sintáticas e semânticas de pronomes pseudoanafóricos no português brasileiro. *Cadernos do IL* 63. 35–61. <https://seer.ufrgs.br/index.php/cadernosdoil/article/view/128495> (11 December, 2025).

Carvalho, Janayna Maria da Rocha, & Ana Regina Calindro. 2018. A unified account for the loss of third person clitics in Brazilian Portuguese. In Danniel da Silva Carvalho, & Dorothy Bezerra Silva de Brito (eds), *Pronomes: Morfossintaxe e semântica*, 91–110. Salvador: EdUFBA.

Chomsky, Noam. 1961. Some methodological remarks on Generative Grammar. *Word* 17(2). 219–239. <https://doi.org/10.1080/00437956.1961.11659755>.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, & Juan Uriagereka (eds), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–156. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in linguistics*, 1–52. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/4056.001.0001>.

Collins, Chris, & Paul M. Postal. 2012. *Imposters: A study of pronominal agreement*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262016889.001.0001>.

Collins, Chris, & Edward Stabler. 2016. A formalization of Minimalist syntax. *Syntax* 19(1). 43–78. <https://doi.org/10.1111/synt.12117>.

Corbett, Greville G. 2006. *Agreement* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Corbett, Greville G. 2015. Hybrid nouns and their complexity. In Jürg Fleischer, Elisabeth Rieken, & Paul Widmer (eds), *Agreement from a diachronic perspective*, 191–214. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110399967-010>.

Costa, João, & Maria Cristina Figueiredo Silva. 2006. Nominal and verbal agreement in Portuguese: An argument for Distributed Morphology. In João Costa, & Maria Cristina

Figueiredo Silva (eds), *Studies on agreement*, 24–46. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/la.86.03cos>.

Costa, João, Denilda Moura, Sandra Maria de Brito Pereira, & Conceição Araújo. 2001. Concordância com a gente: Um problema para a teoria de verificação de traços. In Anabela Gonçalves, & Clara Nunes Correia (eds), *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística (APL)*, 639–656.

Costa, João, & Sandra Pereira. 2005. Phases and autonomous features: A case of mixed agreement in European Portuguese. In Martha McGinnis, & Norvin Richards (eds), *Perspective on Phases* (MITWPL 49), 1–15.

Costa, João, & Sandra Maria de Brito Pereira. 2013. *a gente*: Pronominal status and agreement revisited. *The Linguistic Review* 30(2). 161–184. <https://doi.org/10.1515/tlr-2013-0006>.

Culicover, Peter W., Giuseppe Varaschin, & Susanne Winkler. 2022. The radical unacceptability hypothesis: Accounting for unacceptability without universal constraints. *Languages* 7(96). <https://doi.org/10.3390/languages7020096>.

Cyrino, Sonia Maria Lazzarini. 2003. Para a história do português brasileiro: A presença do objeto nulo e a ausência dos clíticos. *Letras de Hoje* 38(1). 31–47. <https://revistaseletronicas.pucrs.br/fale/article/view/14109> (11 December, 2025).

Davies, Mark. 2016. *Corpus Do Português: Web/Dialects*. Online Corpus. <http://www.corpusdoportugues.org/web-dial/> (30 April, 2025).

Déchaine, Rose-Marie, & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33(3). 409–442. <https://doi.org/10.1162/002438902760168554>.

Den Dikken, Marcel. 2001. Pluringulars, pronouns and quirky agreement. *The Linguistic Review* 18(1). 19–41. <https://doi.org/10.1515/tlir.18.1.19>.

Deng, Chenyuan, Antonio Machicao y Priemer, & Giuseppe Varaschin. 2025. Unifying modifiers, classifiers and demonstratives. In Gabriela Bílbíie, & Gerhard Schaden (eds), *Empirical issues in syntax and semantics: Selected papers from CSSP 2023*, 57–90. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.15450432>.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41(1). 87–100. <https://doi.org/10.1146/annurev-anthro-092611-145828>.

Fasold, Ralph W., & Dennis R. Preston. 2007. The psycholinguistic unity of inherent variability: Old Occam whips out his razor. In Robert Bayley, & Ceil Lucas (eds), *Sociolinguistic variation: Theory, methods, and applications*, 45–69. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511619496.004>.

Fernandes, Eliene. 2004. Fenômeno variável: *nós e a gente*. In Derméval Hora (ed.), *Estudos sociolinguísticos: Perfil de uma comunidade*, 149–162. João Pessoa: ILAPEC.

Freitag, Raquel Meister Ko. 2016. Uso, crença e atitudes na variação na primeira pessoa do plural no português brasileiro. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada* 32(4). 889–917. <https://doi.org/10.1590/0102-44506992907750337>.

Freitag, Raquel Meister Ko, Paloma Batista Cardoso, & Paula Yasmin Gois. 2018. A percepção da variação na primeira pessoa do plural: Efeitos do monitoramento estilístico e urbanização. In Norma da Silva Lopes, Cristina dos Santos Carvalho, & Constância Maria Borges de Souza (eds), *Fala e contexto no português brasileiro: Estudos sobre variação e mudança linguísticas*, 65–84. Salvador: EDUNEB.

Freitas, Maylle Lima, & Hebe Macedo Carvalho. 2020. Quem somos *nós* e quem é *a gente*?: Uma abordagem de avaliação linguística e social da variável de primeira pessoa plural. In Rakel Beserra de Macêdo Viana, Lorena da Silva Rodrigues, Valdecy de Oliveira Pontes, & Hebe Macedo Carvalho (eds), *Estudos em sociolinguística variacionista e sociofuncionalismo*, 124–142. São Paulo: Pimental Cultural. <https://doi.org/10.31560/pimentacultural/2020.985.124-142>.

Green, Georgia M. 1994. The structure of context: The representation of pragmatic restrictions in HPSG. In James H. Yoon (ed.), *Proceedings of the 5th Annual Conference of the Formal Linguistics Society of Mid-America*, vol. 1/2 (Studies in the Linguistic Sciences), 215–232. Urbana, IL: Department of Linguistics, University of Illinois.

Gutzmann, Daniel. 2011. Expressive modifiers & mixed expressives. In Olivier Bonami, & Patricia Cabredo Hofherr (eds), *Empirical issues in syntax and semantics: Selected papers from CSSP 2011*, vol. 8, 123–141. Paris.

Harley, Heidi. 2008. When is a syncretism more than a syncretism? Impoverishment, metasyncretism, and underspecification. In Daniel Harbour, David Adger, & Susana Béjar (eds), *Phi Theory: Phi-features across modules and interfaces*, 251–294. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780199213764.003.0009>.

Haug, Dag Trygve Truslew, & Tatiana Nikitina. 2016. Feature sharing in agreement. *Natural Language & Linguistic Theory* 34. 865–910. <https://doi.org/10.1007/s11049-015-9321-9>.

Heim, Irene, & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Malden, MA: Blackwell.

Kaplan, David. 1999. The meaning of *ouch* and *oops*. Manuscript.

Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In Robert D. Levine, & Georgia M. Green (eds), *Studies in contemporary phrase structure*

grammar, 223–274. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511554421.007>.

Kato, Mary Aizawa, & Maria Eugênia Lammoglia Duarte. 2014. Restrições na distribuição de sujeitos nulos no português brasileiro. *Veredas-Revista de Estudos Linguísticos* 18(1).

Kaur, Gurmeet, & Akitaka Yamada. 2022. Honorific (mis)matches in allocutive languages with a focus on Japanese. *Glossa: A Journal of General Linguistics* 44(1). <https://doi.org/10.16995/glossa.7669>.

Kim, Jong-Bok, & Peter Sells. 2007. Korean honorification: A kind of expressive meaning. *Journal of East Asian Linguistics* 16(4). 303–336. <https://doi.org/10.1007/s10831-007-9014-4>.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237. <https://doi.org/10.1162/ling.2009.40.2.187>.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4). 715–762. <https://doi.org/10.2307/412333>.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Lima, Elyne Giselle Aguiar de Santana. 2015. Variação *nós* e *a gente* na posição de sujeito na escrita escolar. *Letras & Letras* 31(2). 128–143.

Lopes, Célia Regina dos Santos. 1998. *nós* e *a gente* no português falado culto do Brasil. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada* 14. 405–422. <https://doi.org/10.1590/S0102-44501998000200006>.

Lopes, Célia Regina dos Santos. 1999. *A inserção de a gente no quadro pronominal Português: Percurso histórico*. Rio de Janeiro: Universidade Federal do Rio de Janeiro Dissertation.

Lopes, Célia Regina dos Santos. 2004. A gramaticalização de *a gente* em português em tempo real de longa e de curta duração: Retenção e mudança na especificação dos traços intrínsecos. *Fórum linguístico* 4(1). 47–80.

Machicao y Priemer, Antonio, Pauline Reiß, & Giuseppe Varaschin. 2025. Genitive vs. PP arguments in German NPs: Grammatical and use conditions. In Stefan Müller, & Shûichi Yatabe (eds), *Proceedings of the 32nd International Conference on Head-Driven Phrase Structure Grammar, Centro de Linguística da Universidade de Lisboa*. <https://ling.auf.net/lingbuzz/009540> (1 December, 2025).

Maia, Francisca Paula Soares. 2009. A variação *nós/a gente* no dialeto mineiro: Investigando a transição. *Revista da ABRALIN* 8(2). 45–70.

Maia, Francisca Paula Soares. 2012. *Investigando as formas reduzidas de a gente no dialeto mineiro*. Belo Horizonte: Universidade Federal de Minas Gerais Dissertation.

Marcotulio, Leonardo, Juliana Barbosa de Segadas Vianna, & Célia Regina dos Santos Lopes. 2013. Agreement patterns with *a gente* in portuguese. *Journal of Portuguese Linguistics* 12(2). 125–149. <https://doi.org/10.5334/jpl.70>.

Mattos, Shirley Eliany Rocha. 2013. *Goiás na primeira pessoa do plural*. Brasília: Universidade de Brasília Dissertation.

McCready, Elin. 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3. 1–57. <https://doi.org/10.3765/sp.3.8>.

McCready, Elin. 2019. *The semantics and pragmatics of honorification: Register and social meaning*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198821366.001.0001>.

de Mendonça, Alexandre Kronemberger. 2012. *nós e a gente* na cidade de Vitória: Análise da fala capixaba. *PERcursos Linguísticos* 2(4). 1–18. <https://periodicos.ufes.br/percursos/article/view/3173> (11 December, 2025).

Menuzzi, Sérgio. 2000. First person plural anaphora in Brazilian Portuguese: Chains and constraint interaction in binding. In João Costa (ed.), *Portuguese syntax: New comparative studies*, 191–240. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780195125757.003.0008>.

Menuzzi, Sergio. 1995. 1st person plural anaphora in Brazilian Portuguese and chains. *Linguistics in the Netherlands* 12(1). 151–162.

Menuzzi, Sergio, & Carlos Mioto. 2006. Advérbios monossilábicos pós-verbais no PB: sobre a relação entre sintaxe e prosódia. *Revista de Estudos da linguagem* 14(2). 211–243. <https://periodicos.ufmg.br/index.php/relin/article/view/28764> (11 December, 2025).

Menuzzi, Sérgio, & Maria Lobo. 2016. Binding and pronominal forms in Portuguese. In W. Leo Wetzels, João Costa, & Sérgio Menuzzi (eds), *The Handbook of Portuguese linguistics*, 338–355. Malden, MA: John Wiley & Sons. <https://doi.org/10.1002/9781118791844.ch18>.

Meyerhoff, Miriam, & James A. Walker. 2013. An existential problem: The sociolinguistic monitor and variation in existential constructions on Bequia. *Language in Society* 42(4). 407–428. <https://doi.org/10.1017/S0047404513000456>.

Monteiro, Jose Lemos. 1994. *Pronomes pessoais*. Fortaleza: EUFC.

Naro, Anthony Julius, Edair Görski, & Eulália Fernandes. 1999. Change without change. *Language Variation and Change* 11(2). 197–211. <https://doi.org/10.1017/S0954394599112043>.

Nevins, Andrew, & Jeffrey K. Parrott. 2010. Variable rules meet impoverishment theory: Patterns of agreement leveling in English varieties. *Lingua* 120(5). 1135–1159. <https://doi.org/10.1016/j.lingua.2008.05.008>.

de Omena, Nelize Pires. 1996. A referência à primeira pessoa no plural. In Giselle Machline Silva, & Maria Marta Pereira Scherre (eds), *Padrões sociolinguísticos*, 311–323. Rio de Janeiro: Tempo Brasileiro.

Panagiotidis, Phoevos. 2002. *Pronouns, clitics and empty nouns: 'Pronominality' and licensing in syntax*. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/la.46>.

Paolillo, John C. 2000. Formalizing formality: An analysis of register variation in Sinhala. *Journal of Linguistics* 36(2). 215–259. <https://doi.org/10.1017/S002226700008148>.

Pereira, Deize Crespim. 2007. *Variação e mudança no uso dos pronomes reflexivos no português popular da capital paulista*. São Paulo, Brazil: Universidade de São Paulo Dissertation.

Pereira, Sandra Maria de Brito. 2003. *Gramática comparada de a gente: Variação no português europeu*. Lisbon: Universidade de Lisboa Dissertation.

Pesetsky, David, & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian, & Wendy K. Wilkins (eds), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 262–294. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/la.101>.

Podobryaev, Alexander. 2014. *Persons, imposters, and monsters*. Cambridge, MA: Massachusetts Institute of Technology Dissertation.

Pollard, Carl, & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.

Postal, Paul M. 1969. On so-called *pronouns* in English. In David A. Reibel, & Sanford A. Schane (eds), *Modern studies in English: Readings in Transformational Grammar*, 201–224. Englewood Cliffs: Prentice-Hall.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199273829.001.0001>.

Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33(2). 165–198. <https://doi.org/https://doi.org/10.1515/TL.2007.011>.

Potts, Christopher, & Shigeto Kawahara. 2004. Japanese honorifics as emotive definite descriptions. In Robert B. Young (ed.), *Proceedings of the 14th Semantics and Linguistic Theory Conference (SALT)*, vol. 14, 253–270. Northwestern University: The Linguistic Society of America. <https://doi.org/10.3765/salt.v14i0.2917>.

Puškar, Zorica. 2017. *Modelling variation, hierarchy effects and ϕ -feature mismatches*. Universität Leipzig Dissertation.

Ramos, Jania M., & Francisca Paula Soares Maia. 2015. Do estatuto pronominal de *a gente* e cliticização. *Estudos Linguísticos* 44(1). 390–400. <https://revistas.gel.org.br/estudos-linguisticos/article/view/897> (11 December, 2025).

Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.

Robinson, Mary. 2022. *Negative concord as a window into social perception of morphological and syntactic variables*. New York, NY: NYU Dissertation.

Rubio, Cássio Florêncio, & Sebastião Carlos Leite Gonçalves. 2012. A fala do interior paulista no cenário da sociolinguística brasileira: Panorama da concordância verbal e da alternância pronominal. *Alfa: Revista de Linguística* 56(1). 1003–1034. <https://doi.org/10.1590/S1981-57942012000300012>.

Saab, Andrés. 2021. On the locus of expressivity: Deriving parallel meaning dimensions from architectural considerations. In Eleonora Orlando, & Andrés Saab (eds), *Slurs and expressivity: Semantics and beyond*, 17–44. Lanham, MA: Lexington Books. <https://doi.org/10.5040/9781978728059.ch2>.

Sauerland, Uli. 2004a. A comprehensive semantics for agreement. Unpublished Manuscript.

Sauerland, Uli. 2004b. A team, definitely. *Snippets* 9. 11–12.

Schwenter, Scott, Paige Barton, Kendra Dickinson, & Marcia Macedo. 2022. Speakers' subjective evaluations of direct object pronouns in Brazilian Portuguese. *Toronto Working Papers in Linguistics* 44(1). <https://doi.org/10.33137/twpl.v44i1.36730>.

Seara, Izabel Christine. 2000. A variação do sujeito nós e a gente na fala florianopolitana. *Organon* 14(28-29). 179–194. <https://doi.org/10.22456/2238-8915.30203>.

Smith, Peter W. 2017. The syntax of semantic agreement in English. *Journal of Linguistics* 53(4). 823–863. <https://doi.org/10.1017/S002226716000360>.

Smith, Peter W. 2021. *Morphology-semantics mismatches and the nature of grammatical features*. Berlin: De Gruyter. <https://doi.org/10.1515/9781501511127>.

Soares, Eduardo Correa, Philip H. Miller, & Barbara Hemforth. 2020. The effect of semantic and discourse features on the use of null and overt subjects: A quantitative study of third person subjects in Brazilian Portuguese. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada* 36. <https://doi.org/10.1590/1678-460X2020360107>.

Stump, Gregory. 2001. *Inflectional morphology: A theory of paradigm structure*. Vol. 93. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486333>.

Sudo, Yasutada. 2012. *On the semantics of phi features on pronouns*. Cambridge, MA: MIT Dissertation.

Taniguchi, Ai. 2019. Social meaning at the semantics-sociolinguistics interface. Unpublished Manuscript. <https://aitaniguchi.github.io/socialmeaning080119-toshare.pdf> (11 December, 2025).

Taylor, Michael. 2009. On the pronominal status of Brazilian Portuguese *a gente*. *NYU Working Papers in Linguistics* 2. 1–36. https://as.nyu.edu/content/dam/nyu-as/linguistics/documents/nyuwpl/taylor_09_a_gente_nyuwpl2.pdf (22 July, 2023).

Van Eynde, Frank. 2006. NP-internal agreement and the structure of the noun phrase. *Journal of Linguistics* 42(1). 139–186. <https://doi.org/10.1017/S002226705003713>.

Varaschin, Giuseppe. 2021a. *A simpler syntax of anaphora*. Florianopolis: Universidade Federal de Santa Catarina Dissertation.

Varaschin, Giuseppe. 2021b. The disunity of principle B effects. In Stefan Müller, & Nurit Melnik (eds), *Proceedings of the 28th international conference on Head-Driven Phrase Structure Grammar, Frankfurt am Main*, 220–240. Frankfurt am Main: University Library. <https://doi.org/10.21248/hpsg.2021.12>.

Varaschin, Giuseppe, Antonio Machicao y Priemer, & Yanru Lu. 2024. Topic drop in German: Grammar and usage. In Stefan Müller, & Rui P. Chaves (eds), *Proceedings of the 31st International Conference on Head-Driven Phrase Structure Grammar, Palacký University Olomouc, Czech Republic*, 193–217. Frankfurt am Main: University Library. <https://doi.org/10.21248/hpsg.2024.12>.

Varaschin, Giuseppe, Antonio Machicao y Priemer, & Elin McCready. 2025. When I am you: Deriving honorifics through weak indices. Paper presented at NELS 56 [to appear]. <https://box.hu-berlin.de/f/f8365c0cd1434ede8579/> (1 December, 2025).

Vianna, Juliana Barbosa de Segadas. 2006. *A concordância de nós e a gente em estruturas predicativas na fala e na escrita carioca*. Rio de Janeiro: Universidade Federal do Rio de Janeiro MA thesis.

Vianna, Juliana Barbosa de Segadas, & Célia Regina dos Santos Lopes. 2012. A variação entre *nós* e *a gente*: Uma comparação entre o português europeu e o português brasileiro.

Revista do GELNE 14(1/2). 95–116. <https://periodicos.ufrn.br/gelne/article/view/9366> (11 December, 2025).

Vianna, Juliana Barbosa de Segadas, & Célia Regina dos Santos Lopes. 2015. Variação dos pronomes *nós* e *a gente*. In Marco Antonio Martins, & Jussara Abraçado (eds), *Mapamento sociolinguístico do português brasileiro*, 109–131. São Paulo: Contexto.

Wechsler, Stephen. 2021. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley, & Jean-Pierre Koenig (eds), *Head-Driven Phrase Structure Grammar: The Handbook*, 217–242. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.5599828>.

Wechsler, Stephen, & Larisa Zlatić. 2003. *The many faces of agreement*. Stanford: Center for the Study of Language & Information.

Wurmbrand, Susi. 2017. Formal and semantic agreement in syntax: A dual feature approach. In *Language use and linguistic structure: Proceedings of the Olomouc Linguistics Colloquium 2016*, 19–36. Olomouc: Palacký University.

Zilles, Ana M. S. 2002. Grammaticalization of *a gente* in Brazilian Portuguese. *University of Pennsylvania Working Papers in Linguistics* 8(3). 297–310. <https://repository.upenn.edu/handle/20.500.14332/45604> (11 December, 2025).

Zilles, Ana M. S. 2005. The development of a new pronoun: The linguistic and social embedding of *a gente* in Brazilian Portuguese. *Language Variation and Change* 17(1). 19–53. <https://doi.org/10.1017/S0954394505050027>.

Zilles, Ana M. S. 2007. O que a fala e a escrita nos dizem sobre a avaliação social do uso de *a gente*? *Letras de Hoje* 42(2). 27–44. <https://revistaseletronicas.pucrs.br/fale/article/view/2408> (11 December, 2025).