'Lexical DP Blocking' in the Person-Case Constraint as a licensing failure: Comparative evidence from Mixtec

1 Introduction

Many languages display person-based restrictions on combinations of two pronominal clitics, a phenomenon now known as the Person-Case Constraint (PCC) (Perlmutter, 1971; Bonet, 1991, et seq.). A common approach to the PCC is that it arises from a single probe attempting to—but failing to—Agree with both clitics, though individual analyses vary in the exact mechanics of Agree, as well as the loci and feature specifications of the probe and goals.[1] Sichel and Toosarvandani (to appear) show that, in Southeastern Sierra Zapotec (henceforth, 'SSZ'), an Eastern Otomanguean language group, the PCC restricts combinations of monotransitive external and internal arguments (henceforth, subjects and objects).[2] As illustrated in (1a-b) with the Laxopa variety,[3] SSZ displays the *strong PCC*, in that 1st and 2nd person but not 3rd person object clitics are banned in the context of a higher subject clitic. Importantly, these authors also observe that the SSZ displays another PCC-like pattern, otherwise unattested in the typology of PCC effects: object clitics of *all persons* (i.e., 1, 2, *and* 3) are banned when the subject is a lexical DP rather than a pronominal clitic, (1c). Sichel and Toosarvandani frame the latter as a *Lexical DP Blocking effect (LDB)*: the subject DP intervenes between a higher Agreeing probe P and the object pronoun, preventing it from cliticizing to the verb, (2). Moreover, lexical DPs and pronouns may interact

---

with P in turn because they share featural content, both bearing (at least) a D-feature $[\delta]$.[4]

(1) a. *Ba   betw=**ba'=o'**
       already CMPLL.hit=3.HU=2SG
       Intended: 'S/he already hit you.'
       (Laxopa; toosarvandani2017, p. 128)

    b. Blenh^3=**ba'^{2}=b^3**
       CMPLL.carry=3.HU=3.AN
       'S/he carried it.'
       (Laxopa; Sichel and Toosarvandani to appear)

    c. *Blenh^3=**eb^3**$_i$ Xwanh^1=a'^{3} t_i
       CMPLL.carry=3.AN Juana=DEF
       Intended: 'Juana carried it.'
       (Laxopa; Sichel and Toosarvandani to appear)

(2)
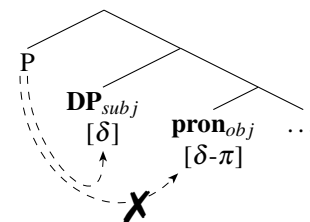
This paper develops an alternative account of the LDB, recasting it as a Case Filter violation that interacts with the Agree mechanism responsible for the PCC. Under this alternative, the LDB arises because the same person-sensitive probe is also independently needed for DP-licensing. Evidence for this alternative comes from the San Juan Piñas variety of Mixtec (SJPM), also Eastern Otomanguean, whose subject and object clitics also display both the PCC and the LDB. Crucially, however, certain DP/pronoun combinations, while wholly ruled out in SSZ, are well-formed in SJPM, (3).[5] These examples also reveal that pronominal object clitics do not move past DP subjects, as in (1c), but simply encliticize to them. This paper thus suggests that, in both Zapotec and Mixtec, the probe responsible for PCC effects, P in (2), is actually structurally low, located between the subject and object. This probe, which I locate in Voice, cyclically expands in the sense of Béjar and Rezac 2009, encountering the object on a first cycle of Agree before then attempting to target the subject, in Spec-VoiceP, on the second cycle, (4).

---

[4]**Abbreviations for SSZ and SJPM data:** AN = animate, BASE = pronominal base, CL = classifier, CMPL = completive, CONT = continuative, DEF = definite, EL = elder, EX = exclusive, F = feminine, HAB = habitual, HON = honorific, HU = human, IRR = irrealis, M = masculine, N = neuter, NEG = negation, PL = plural, SG = singular, 1/2/3 = 1st/2nd/3rd person.

[5]The SJPM data are represented in IPA, pending a conventionalized orthography. Tones are represented using the Chao numeric system. In contrast to the SSZ tonal conventions described in fn. 3, for the SJPM data, V^1 indicates the lowest tone and V^5 is the highest tone (a contextually restricted process of upstep may also produce super-high tones, transcribed as V^6). See Caballero, Duarte Borquez, Juárez Chávez, and Yuan to appear to appear for an overview of the tone system of SJPM.

(3) a. ko^5ni^1 [ti^5 vi^3lu^5]=**ra^3**
         CONT.love   CL.3.AN cat   =3SG.M
         'The cat loves him.'

   b. *ko^5ni^1 [ti^5 vi^3lu^5]=**^{n}di^3**
         CONT.love   CL.3.AN cat   =1PL.EX
         Intended: 'The cat loves us (excl.).'

(4)

$\mathbf{DP}_{subj}$ ② Voice $\mathbf{pron}_{obj}$ ① …
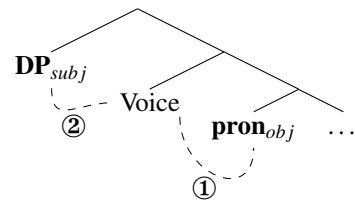
Importantly, the DP subject *never* prevents Voice from Agreeing with the object pronoun—so the so-called LDB does not truly involve Lexical DP Blocking. Instead, building on the theory of PCC effects developed in Deal 2015, 2024, the PCC arises whenever a *pronominal object* prevents second-cycle Agree with the subject, not the other way around. The LDB, in turn, reflects configurations in which a DP subject is left unlicensed by Agree by the end of the derivation, violating the Case Filter. Not only is the probe in Voice unable to target the subject in these configurations, but no other heads may target the subject either.

The latter point, in turn, is related to the *verb-initial syntax* of SSZ and SJPM: a single VoiceP-internal probe is responsible for both PCC effects and subject licensing, as other heads canonically implicated in argument licensing (e.g., T) are either inactive or solely responsible for verb-raising (e.g. Alexiadou and Anagnostopoulou, 1998). However, Voice may only license the subject if second-cycle Agree is not blocked by the PCC. Thus, this paper ultimately reframes the LDB as the convergence of otherwise independent restrictions: in SSZ and SJPM, both the PCC and the Case Filter are calculated VoiceP-internally and thus restrict subjects (pronominal or lexical DP) in Spec-VoiceP in comparable ways. The highly specific syntax invoked in this analysis may also explain the rarity of the LDB cross-linguistically. That being said, the existence of the LDB in both Zapotec and Mixtec varieties suggests that it may hold in Eastern Otomanguean more generally, so this analysis may serve as a blueprint for future investigations in related languages.

This paper is organized as follows. §2 presents the PCC effects found in SSZ and outlines Sichel and Toosarvandani's (to appear) analysis of the LDB. §3, turns to the PCC

and LDB in SJPM and show how these patterns are difficult to capture in terms of subject intervention; §4 pursues an alternative account building on Deal 2024 and Deal and Royer to appear. Finally, §5 argues that the LDB arises in these languages due to their common verb-initial syntax, which delimits the functional heads available for argument licensing.

2   The PCC and LDB in Southeastern Sierra Zapotec

2.1   Overview of morphosyntactic patterns

Zapotec is part of the Eastern Otomanguean subgroup of the larger Otomanguean language family (Kaufman, 1988; Campbell, 2017a,b). The varieties discussed by Sichel and Toosarvandani (to appear) represent multiple Northern Zapotec varieties from the southeastern Sierra Norte region of Oaxaca, Mexico. The basic facts presented in this section hold for all varieties surveyed by these authors (collectively grouped under Southeastern Sierra Zapotec, or 'SSZ').[6] As we saw in (1), SSZ has a base V(erb)-S(ubject)-O(bject) word order. Following Lee 2006 and Adler et al. 2018, the basic derivation of verb-initial order in these languages involves object shift out of the verb phrase (to an A-position below the external argument), followed by remnant movement of the verbal predicate. In §5, it will be argued that there is no derived A-position for subjects; thus, the term 'subject' refers here to the highest DP within the A-domain, typically the external argument.

(5)
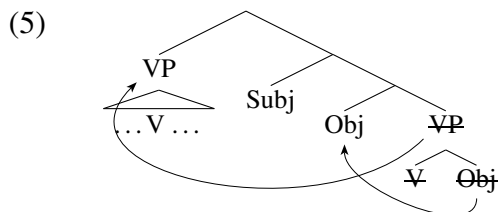
Pronominal subjects and objects in SSZ may be expressed as enclitics, which also surface in VSO order. These clitics are taken by Sichel and Toosarvandani (to appear) to be gen-

---

[6]Sichel and Toosarvandani (to appear) consider the varieties spoken in the towns of Santiago Laxopa, Santa María Yalina, and San Sebastián Guiloxi, with data coming from their own fieldwork, as well as the varieties spoken in Hidalgo Yalálag (citing Avelino Becerra 2004 and López and Newberg 2005) and San Bartolomé Zoogocho (mostly building on Sonnenschein 2004).

erated via Agree (following Sportiche 1993, Anagnostopoulou 2003, a.m.o.); specifically, cliticization results from *pronoun movement*, triggered by the Agreeing probe.

Recall from (1) that SSZ displays the *strong PCC*, summarized in (6). In SSZ, this is a ban on 1st and 2nd person clitics in object position, regardless of the person of the subject. More examples are provided in (7); these data also show that, to express the PCC-violating configurations, the lower argument must instead be encoded as a tonic pronoun.

(6)   *The Strong PCC:* In combinations of two clitics, Cl_1 and Cl_2, where Cl_1 is structurally higher than Cl_2, Cl_2 must be 3rd person.

(7)   a.   Bi   llre'=**o'** |nada'|
            NEG HAB.see=2SG 1SG
            'You don't see me.' (*1SG =*a*')

      b.   Bi   llre'=**la'** |lue'|
            NEG HAB.see=1SG 2SG
            'I don't see you.' (*2SG =*o*')

      c.   Bet=gak=**a'**=**ba'**
            CMPL.kill=PL=1SG=3.AN
            'I killed them (animals).'
            (Yalálag;   Foley   and   Toosarvandani 2022)

Furthermore, in combinations of 3rd person subject and object clitics, the object clitic generally cannot outrank the subject along an animacy hierarchy (Foley et al., 2019; Foley and Toosarvandani, 2022), a pattern termed the *Gender-Case Constraint* (on par with the PCC). The GCC will be discussed in §4.1. Finally, no object clitic *of any person* is permitted if the subject is a lexical DP. This was seen in (1c) and further illustrated in (8). Sichel and Toosarvandani (to appear) frame this as a *Lexical DP Blocking* effect (LDB), the intuition being that the presence of a lexical DP blocks the object from cliticizing. They further claim that the LDB is a novel pattern within the existing typology of PCC patterns, though observe that it is akin to the *weak PCC*: if a monotransitive contains a lexical DP argument, it must be in subject position.[7] As additional support for a unified approach to the PCC, the GCC, and LDB, they note that the choice of the so-called 'repair' is identical in all cases:

---

[7]The weak PCC canonically states that, if there is one 3rd person clitic, it must be in direct object position (thus, like the strong variant, 3/PART combinations are ruled out, but, unlike the strong variant, PART/PART combinations are licit).

whenever cliticization is blocked, the object is encoded as a tonic pronoun.

(8)   a.   Dz^{3}la^{3}lle'^{3} **Xwanh^1=a'^{3}** $\boxed{\textbf{lhe'}^1}$ b.   Blenh^3 **Xwanh^1=a'^{3}** $\boxed{\textbf{leb}^{13}}$
           CONT.forget Juana=DEF   2SG   CMPLL.carry Juana=DEF   3.AN
           'Juana forgot you.' (*2SG =u'^{3})   'Juana carried it.' (*3.AN =eb^3)
           (Laxopa; Sichel and Toosarvandani to appear)

The LDB is not alleviated when the DP subject is displaced (e.g., focus-fronted) to a pre-verbal position, (9). This reinforces that the LDB is not a purely morphological or prosodic restriction. Rather, the underlying explanation must be structural in nature: in (9), it is the (non-overt) base copy of the DP subject that triggers the LDB.

(9)   **Xwanh^1=a'^{3}**$_i$ blenh^3 t_i $\boxed{\textbf{leb}^{13}}$
       Juana=DEF   CMPLL.carry   3.AN
       'JUANA carried it.' (*3.AN =eb^3)   (Laxopa; Sichel and Toosarvandani to appear)

The PCC and LDB patterns in SSZ may thus be summed up as (10):

(10)   *The Strong PCC/LDB in SSZ (where 1/2/3 are pronominal clitics):*

Subj.	2	1	3	3	DP	DP	DP	3	2	1
Obj.	1	2	1	2	1	2	3	3	3	3
	✗	✗	✗	✗	✗	✗	✗	(✓)[8]	✓	✓

2.2   Subject intervention and a 'Probe Activation' account

As pronominal cliticization is preconditioned on Agree, Sichel and Toosarvandani (to appear) take well-formed configurations containing two pronominal clitics to reflect multiple successful instances of Agree. Violations of the PCC in SSZ would, in turn, result from a probe's failure to Agree with an object clitic. Moreover, they extend this logic to the LDB: a lexical DP subject may also prevent a probe from Agreeing with a lower object clitic.

Much prior work on person hierarchies (subsuming PCC effects) appeal to the featural representations of different person features—with 1st/2nd (i.e., PART(icipant)) person involving a more articulated feature geometry than 3rd person (Harley and Ritter, 2002;

---

[8]As noted above, the well-formedness of individual 3/3 combinations is subject to the GCC.

6

Béjar, 2003, a.o.). To explain why lexical DP subjects also participate in feature hierarchy effects, Sichel and Toosarvandani reason that all nominals of category D—pronouns and DPs alike—bear a dedicated D-feature [δ], as summarized in (11). Pronouns bear an additional feature [π] (cf. Sichel and Wiltschko, 2021); 1st/2nd person pronouns additionally bear [PART]. The feature geometries for full DPs, 3rd person pronouns, and PART pronouns are thus presented below, with features on the right entailing those on the left.

(11)   *Extended Person (Sichel and Toosarvandani, to appear):*
         Lexical DPs share a feature (δ) that is part of the structure of Person; only personal pronouns are specified for additional Person features.

    a. *Lexical DP:*   b. *3rd pers. pronoun:* c. *1st/2nd pers. pronoun*:
       [δ]   [δ-π]   [δ-π-PART]

Sichel and Toosarvandani propose that a single high probe, P, is always able to target the subject, the closest goal in P's c-command domain. However, P may also Agree with the object only if it is *not* more featurally specified than the subject (cf. Nevins 2007). This is formulated as *Probe Activation*, synthesized in (12). Thus, whereas the initial round of Agree takes place to value P, subsequent rounds of Agree are *Greed-based*, in that they take place solely to satisfy requirements of the lower goal G2 (since G2 does not bear additional features not already found on G1). In particular, this is needed to generate object clitics.

(12)   *Probe Activation (adapted from Sichel and Toosarvandani to appear):* After Agreeing with its highest goal G1, the probe is "activated" and is able to Agree with another goal G2, just in case G2 is not featurally distinct from the probe P: i.e., G1 is not more featurally specified than P for the features P searched for (G2 $\subseteq$ P).

Putting this together with the Extended Person component of their proposal, Sichel and Toosarvandani are able to capture the various PCC-like restrictions in SSZ. For instance, consider the LDB, depicted in (13a). If the subject is a lexical DP (bearing only [δ]) and the object is a pronoun (of any person), Probe Activation is impossible since the latter bears (at least) [δ-π]. As such, the object cannot be realized as a clitic (it can, however,

be exponed as a tonic pronoun). The structure in (13b) illustrates the strong PCC, in that P may generate a 3rd person ($[\delta\text{-}\pi]$) subject clitic via Agree, but not a participant ($[\delta\text{-}\pi\text{-}$PART$]$) object clitic. Finally, (13c) illustrates a well-formed two-clitic construction: here, the 3rd person ($[\delta\text{-}\pi]$) pronoun does not bear additional features not already copied by the subject ($[\delta\text{-}\pi\text{-}$PART$]$), so an object clitic may be generated by Probe Activation.

(13)   a. *DP/pron: b. *3/PART: c. PART/3:

This system also extends straightforwardly to the GCC: 3rd person elements higher in animacy are featurally more specified than those lower in animacy, so Probe Activation does not take place if the object outranks the subject.[9]

    In sum, Sichel and Toosarvandani (to appear) take the PCC and LDB in SSZ to arise from three interacting factors: (i) a high probe, c-commanding both arguments of the verb, (ii) an Agree requirement on pronominal cliticization, and (iii) a model of Agree permitting subsequent rounds of Agree only if no new features would be copied to the probe.

3   The PCC and LDB in San Juan Piñas Mixtec

We now turn to novel data from the San Juan Piñas variety of Mixtec (SJPM) that bear on this discussion. SJPM displays both the PCC and (a version of) the LDB, but the exact

---

[9]Lastly, to capture ill-formed *PART/PART combinations also characteristic of the strong PCC (e.g., *1/2 and *2/1), Sichel and Toosarvandani appeal to additional features such as [SPKR] that differentiate 1st from 2nd person. To be precise, 1st and 2nd person are specified as $[\delta\text{-}\pi\text{-}$PART-SPKR$]$ and $[\delta\text{-}\pi\text{-}$PART$]$, respectively. 2/1 configurations thus are ruled out the same way as 3/PART, as in (13b), with the conditions for Probe Activation not being met. However, *1/2 configurations require an additional stipulation: the [PART-SPKR] portion of the feature geometry of 1st person is copied back to the probe as a complex unit. This means that [PART-SPKR] (in the representation of 1st person) and [PART] (in the representation of 2nd person) are not only featurally distinct, but the latter is technically *not* a proper subset of the former. As such, Sichel and Toosarvandani take Probe Activation to not be possible in such configurations either.

patterns that arise are difficult to capture under the Probe Activation account outlined above.

3.1 Language background and key morphosyntactic facts

Like Zapotec, Mixtec an Eastern Otomanguean language. Though Zapotec and Mixtec are categorized in distinct subgroups within Eastern Otomanguean, they share several morphosyntactic similarities relevant to this paper. SJPM is a Southern Baja variety of Mixtec (Josserand 1983) and is spoken in the town of San Juan Piñas, Santiago Juxtlahuaca, Oaxaca, as well as in diaspora communities in the US. This paper comes out of a broader collaborative project documenting, analyzing, and creating linguistic resources for this otherwise undescribed Mixtec variety, as overseen by the second author of this paper, an L1 speaker of SJPM who was born and raised in San Juan Piñas but is now based in California. The data in this paper were obtained through targeted elicitation with the second author, though the key empirical patterns were additionally confirmed with three other speakers of SJPM. The paper's analytical and theoretical claims were developed by the first author.[10]

SJPM also displays base V(P)-S-O word order, as in (5), and restricts certain combinations of subject and object enclitics. In (14a-b), we see that SJPM displays the strong PCC, in that 1st and 2nd person (PART) objects must surface as tonic pronouns, not clitics; in contrast, there is no restriction on 3rd person object clitics, (14c).[11] There are no animacy-based constraints on combinations of 3rd person subject and object clitics (§4.1).

(14) a. $ʃi^{13}ni^{31}$=$^ndo^5$ $\boxed{^ndu^1(ʔu^1)}$ b. ko^5ni^1=$ɲa^5$ $\boxed{^ndu^1(ʔu^1)}$
        CMPL.see=2PL 1PL.EX   CONT.love=3SG.F 1PL.EX
        'You (pl.) saw us (ex.).'   'She loves us (excl.).'
        (*1PL.EX =$^ndi^3$)   (*1PL.EX =$^ndi^3$)

---

[10] The pronoun 'I' is used (especially in §§3.2-5) when referring to the theoretical proposals of the first author.

[11] SJPM displays a three-way contrast between (monomoraic) pronominal clitics, (monomoraic) weak pronouns, and (bimoraic) strong pronouns. Importantly, both weak and strong pronouns are considered in this paper to be 'tonic pronouns'. As discussed by Macaulay (1987) for Chalcatongo Mixtec, weak pronouns are effectively reduced versions of the strong pronouns (e.g., used in faster speech) and thus pattern *morphosyntactically* like strong pronouns, even though they are monomoraic like the clitics. Indeed, as shown throughout this paper, both weak and strong PART pronouns may appear in object position in SJPM, not restricted by the PCC, suggesting that they form a natural class to the exclusion of the pronominal clitics.

c. ko^{5}ni^1=^{n}**di^1**=**na^1**
CONT.love=1PL.EX=3PL.N
'We (excl.) love them.'

Importantly, the LDB holds in SJPM as well.[12] The examples in (15) mirror those from SSZ above, in that the LDB is not alleviated by Ā-extracting the lexical DP subject (here, a fronted negative quantifier) past the verb.

(15) a. ko^{5}ni^1 [ti^5 vi^{3}lu^5] $\boxed{^n\textbf{du}^1\textbf{(}^?\textbf{u}^1\textbf{)}}$
CONT.love   CL.3.AN cat   1PL.EX
'The cat loves us (excl.).' (*1PL.EX =^{n}di^3)

b. **ni^3=ñ^3=na^1**$_i$ ko^{15}-ʃi^{3}ni^{31} t_i $\boxed{^n\textbf{du}^1\textbf{(}^?\textbf{u}^1\textbf{)}}$
NEG=one=3PL.N NEG.CMPL-see   1PL.EX
'No one saw us (excl.).' (*1PL.EX =^{n}di^3)

However, the exact pattern found in SJPM diverges from that in SSZ in one crucial respect: the LDB only affects PART objects. In contrast to (15), the examples in (16) show that 3rd person pronominal objects are licit as clitics in the same environments—they simply encliticize to the immediately preceding element (a fact that will be important to the analysis). In verb-initial word order, the clitic host is the postverbal subject, but, if the subject is Ā-extracted, then the object clitic simply attaches to the verb.

(16) a. ko^{5}ni^1 [**ti^5**   **vi^{3}lu^5**]=**ɾa^3**
CONT.love   CL.3.AN cat   =3SG.M
'The cat loves him.'

b. **ni^3=ĩ^{3}ĩ^3=na^1**   ko^{15}-ʃi^{3}ni^{31}=**na^1**
NEG=one=3.PL NEG.CMPLL-see=3.PL
'No one saw them.'   ni^3=ĩ^{3}ĩ^3=na^1 ko^{15}-ʃi^{3}ni^{31} t_i =na^1

---

[12] While this paper focuses mainly on SJPM, the inability for 1st/2nd person objects to be expressed as clitics (regardless of the status of the subject) is a broader property of Mixtec, having been cursorily mentioned in various descriptive works for at least Chalcatongo Mixtec (Macaulay, 1996, pp. 140-141), Alacatlatzala Mixtec (Zylstra, 2012, p. 64), Alcozauca Mixtec (Cline, 2018, p. 24), Ixtayutla Mixtec (Penner, 2019, pp. 26, 52-54), and Ocotepec Mixtec (Alexander 1988, pp. 263-265, Lopez 2022, pp. 54-57). None of these descriptions explicitly frame such restrictions as the PCC and there are no prior analyses.

The PCC and LDB in SJPM is summarized in (17), with the crucial divergence from the SSZ pattern boxed (cf. (10)); additional differences will be presented later in this paper. Notably, the LDB in SJPM is an extension of the strong PCC: [PART] object clitics are ruled out regardless of the choice of subject.

(17) *The Strong PCC/LDB in SJPM (where 1, 2, and 3 are pronominal clitics):*

Subj.	2	1	3	3	DP	DP	DP	3	2	1
Obj.	1	2	1	2	1	2	3	3	3	3
	✗	✗	✗	✗	✗	✗	☑	✓	✓	✓

Given Sichel and Toosarvandani's (to appear) claim that the LDB is otherwise unattested within the empirical landscape of the PCC, its existence in SJPM is notable. As Mixtec and Zapotec are not closely related, categorized within distinct subbranches of Eastern Otomanguean, this suggests that the LDB may be a general property of this language group—a testable prediction that can be confirmed in future work.[13] Yet, the morphosyntactic similarities between SSZ and SJPM suggest that the differences pertaining to the PCC and LDB should be modeled as parameterization within a common grammatical structure.

3.2 The LDB without subject intervention

To extend Sichel and Toosarvandani's (to appear) account to SJPM, we should use the featural representations from their Extended Person hypothesis, as in (11). However, the exact patterns seen in SJPM present empirical challenges to Sichel and Toosarvandani's Probe Activation analysis. First, a lexical DP ($[\delta]$) subject should prevent the higher probe, P, from targeting 1st/2nd person ($[\delta\text{-}\pi\text{-PART}]$) and 3rd person ($[\delta\text{-}\pi]$) pronominal object clitics alike, since pronominal clitics are uniformly more featurally specified than lexical DPs. While this is true for SSZ, repeated in (18a), recall that DP/3 combinations are in

---

[13]In addition, the SJPM LDB pattern is highly reminiscent of Postal's (1989) 'Fancy Constraint' found in *faire-infinitif* causative constructions in certain Romance languages. In such constructions, lexical DP causees may participate in the strong PCC, preventing [PART] but not 3rd person direct object clitics. The SJPM data thus allow us to draw a novel empirical connection between the LDB and the Fancy Constraint (not noted in Sichel and Toosarvandani to appear). See §5.2 for discussion.

fact licit in SJPM. The fact that SJPM bans *DP/PART combinations yet permits DP/3 combinations, as shown in (18b), is contradictory under this account.

(18) a. *LDB in SSZ:*   b. *LDB in SJPM:*

```
              P                                        P
         ╱    ╲                                   ╱    ╲
      Subj     ╲                                Subj    ╲
      ✓[δ]    Obj    . . .                      ✓[δ]   Obj    . . .
             ✗[δ-π]                                   ✓[δ-π]
           ✗[δ-π-PART]                              ✗[δ-π-PART]
```

A related issue arises from Sichel and Toosarvandani's claim that object clitics are blocked from moving past a lexical DP subject. In actuality, SSZ does not offer conclusive evidence either way about the locus of pronominal cliticization in such contexts: object clitics are banned both before and after lexical DP subjects, (19a-b), and combinations of subject/object clitics are uninformative since both arguments cliticize to the same host, (19c).

(19) a. *Blenh³=**eb³**   Xwanh¹=a'³   b. *Bdel   Maria=**b**
          CMPL.carry=3.AN Juana=DEF   CMPL.hug Maria=3.AN
          Intended: 'Juana carried it.'   Intended: 'Maria hugged it.'
          (Laxopa; Sichel and Toosarvandani to   (Sichel and Toosarvandani, 2020)
          appear)

      c. Blenh³=**ba'²**=**b³**
          CMPL.bite=3.HU=3.AN
          'S/he carried it (an animal).'   (Laxopa; Sichel and Toosarvandani to appear)

In earlier work, Sichel and Toosarvandani (2020) provide a possible exception to the LDB: 1st/2nd person pronouns in narrow focus contexts are realized as tonic pronouns and must be clitic-doubled (dashed in (20a)). For transitive constructions containing both a focused clitic-doubled subject and a pronominal object, the latter attaches to the clause-initial verb—ostensibly raising past the (in situ) tonic pronoun subject, (20b).

(20) a. Yega'an=**o'** **lhe'**   b. Betw=**a'**=**ba'**ᵢ   **neda'** tᵢ
          IRR.stay=2SG 2SG   CMPL.hit=1SG=3.HU 1SG
          'You are going to stay.'   'I hit her/him.'
                                            (Sichel and Toosarvandani, 2020)

However, it is not immediately obvious how clitic-doubling the subject allows the object to bypass the LDB: even though clitic doubling is often analyzed as involving a movement dependency (e.g. Anagnostopoulou, 2003; Harizanov, 2014; Baker and Kramer, 2018), we have already seen in examples such as (9) that subject movement otherwise does not alleviate the LDB. The pattern in (20b) is moreover complicated by the fact that object clitics may not raise past focused *3rd person* tonic subjects (which may not be clitic-doubled for independent reasons), (21), a point acknowledged in Sichel and Toosarvandani 2020.

(21)   *Blhe'e=**b**$_i$ **leba'** t_i
         CMPL.see=3.AN 3.HU
         Intended: '<u>S/he</u> saw it.'   (Sichel and Toosarvandani, 2020)

In the absence of additional adjudicating data, it is worth noting that this pattern is compatible with a range of alternative analyses—so it may not tell us anything about the LDB at all. Suppose, for instance, that (20b) actually involves something like *Clitic Right Dislocation*, with the focused pronoun occupying a right-peripheral clausal position.[14] Under this alternative parse, schematized in (23) below, the morpheme order in (20b) would be uninformative about the syntactic height of the object clitic.

(22)   [[VERB=**Cl**$_i$=**Cl**$_j$ t_i t_j] **PART.PRON**$_i$]

Crucially, the locus of object cliticization *is* evident in SJPM, since DP/3 combinations are well-formed. As already seen in (16a), object clitics simply attach to the *immediately preceding subject*; the opposite order is unattested, (23a). This is further exemplified in (23b). This example, which features a focused tonic pronoun subject (without clitic-doubling), offers an important counterpoint to the derivation sketched for (20b) above: the pronominal object clitic does not raise past the focused subject but again encliticizes to it.[15]

---

[14]To rule out (21), we might add that focused 3rd person pronouns cannot undergo CLRD. Though this is a stipulation, the point is that the contrast between (20b) and (21) lacks explanation under *any* account.

[15]Lastly, this example provides phonological evidence for the clitic status of the post-subject element. In SJPM, certain enclitics that otherwise surface as [M] (V^3) may serve as targets of a word-internal L-tone (V^1)

(23) a. *ko⁵ni¹=ɾa¹ᵢ [ti⁵ vi³lu⁵] tᵢ
CONT.love=3SG.M CL.3.AN cat
Intended: 'The cat loves him.'

b. no³mi³ ⁿdu¹ʔu¹=ɾa¹
IRR.hug 1PL.EX=3SG.M
'We (ex.) will hug him.'

As mentioned above, this paper aims to develop a unified analysis of SSZ and SJPM, given their similarities. However, Sichel and Toosarvandani's model cannot easily account for SJPM. Therefore, I instead pursue a *low probe* alternative, wherein P is situated *between* the subject and object. This is supported by the position of object clitics in SJPM and is compatible with available SSZ data as well. Thus, the probe selects for the subject as its specifier and is able to undergo multiple Agree with both arguments via *cyclic expansion* (Béjar and Rezac, 2009): P first Agrees with the object in its c-command domain, (24a), before targeting the subject in a second cycle of Agree, (24b).

(24) a.

P
Obj …

b.

Subj
P
Obj …

Notably, this alternative account maintains the assumption that pronominal clitics are generated by Agree with P, but the ancillary claim that they undergo movement to P is no longer necessary: if movement does take place, it would have to be string-vacuous.[16]

spreading process stemming from the final tone bearing unit of the stem (Duarte Borquez, 2022). In (23), the same 3SG.M clitic is attached to a L-final stem and is realized as [L].

[16]Foley and Toosarvandani (2022) cite the inability for pronominal clitics to be formed in coordinate structures as evidence for movement, (i), under the interpretation that this would involve CSC island violations. The same facts hold in SJPM, but, since SJPM makes a three-way distinction between tonic pronouns, weak pronouns, and clitics, we can see in (ii) that *both* clitics and weak pronouns (reduced versions of tonic pronouns) are banned in coordinate structures, even though the latter are not generated by Agree and thus do not undergo movement (see fn. 11). Instead, such restrictions could be attributed to prosodic or morphophonological requirements on coordinands, orthogonal to movement (e.g. Cardinaletti and Starke, 1999).

(i) Ts-ja-wia [le' na xna'=a] taw=a'
CONT-AND-visit 3.EL and mother=1SG grandmother=1SG
'S/he and my mother went to visit my grandmother.' (*3.EL =e') (Laxopa; Foley and Toosarvandani 2022)

(ii) ka³ta³ [ʒu¹ʔu¹ a⁵ ʐo⁵ʔo¹]
IRR.sing 1SG or 2SG
'You or I will sing.' (*1SG =e¹, *1SG ʒu¹, *2SG =o⁵, *2SG ʐo⁵¹) (SJPM)

But if P finds the object first, then it becomes impossible to characterize the LDB in terms of *subject intervention* (in other words, the term 'Lexical DP Blocking' is a misnomer). As schematized in (24), P never encounters the subject before the object; the subject is only targeted under second-cycle Agree *after* P seeks the object. So what rules out the ill-formed configurations containing lexical DP subjects? I propose that the LDB actually arises from the features of the *object* blocking Agree with the subject.

4   A low probe account

While many analyses of the PCC take a single probe P to c-command both clitics (possible goals for P), low probe accounts of the PCC (involving cyclic expansion of the probe) have also been developed (e.g., Pancheva and Zubizarreta, 2018; Deal, 2024).[17] This paper adopts the approach developed in Deal 2024, based on Deal's (2015) *interaction/satisfaction* theory of Agree. This is because it is specifically cited by Sichel and Toosarvandani (to appear) as non-viable for SSZ due to the LDB; unlike pronominal clitics, lexical DPs are not generated by Agree. We will see that this theory *is* in fact compatible with the LDB—when coupled with additional constraints on DPs such as the Case Filter.

4.1   *Modeling the PCC and LDB using interaction/satisfaction*

Under the theory of Agree developed in Deal 2015, probes are specified for *interaction* and *satisfaction* conditions (henceforth, 'int/sat'), dictating which features are copied back to a probe P, vs. which ones halt further probing by P. This departs from a more traditional conception of Agree as eliminating P's uninterpretable or unvalued features (Chomsky, 2000, 2001). Deal 2024 extends this theory to the PCC: the PCC arises when the first goal that P encounters bears features that match P's satisfaction condition, preventing P from

---

[17]To be precise, in Deal 2024, it is only important that P targets the direct object prior to targeting the other argument (whether the subject of a monotransitive verb or indirect object of a ditransitive verb). Thus, a high probe structure is also compatible with Deal's account, if the object raises past the other argument. This option will not be further considered here, since there is no evidence for this inversion in SJPM or SSZ.

interacting with a second potential goal. For instance, Deal takes the strong PCC to arise when the probe P bears [INT:π,SAT:PART]: P may copy any [π] it encounters, but probing is halted if it finds a goal bearing [PART]. For PCC effects in monotransitive configurations, P is satisfied by a [PART] object, so it does not further probe for the subject, P's specifier.

This logic accounts for strong PCC patterns among subject and object clitics, but it requires a small amendment to handle the LDB. Since lexical DPs bear [δ] but not [π] (Sichel and Toosarvandani, to appear), let us assume that, in SJPM and SSZ, P bears [INT:δ,SAT:PART]. The trees in (25) illustrate how this system applies to SJPM; recall that the LDB in SJPM is essentially the strong PCC. If the object only bears [δ-π] or just [δ], P may continue to probe for the subject (whether pronominal or lexical DP), (25a); however, if the object bears [PART], P is satisfied and does not probe for the subject, (25b). Why non-agreement with the subject ultimately crashes the derivation is addressed in §5.

(25) a.

b.
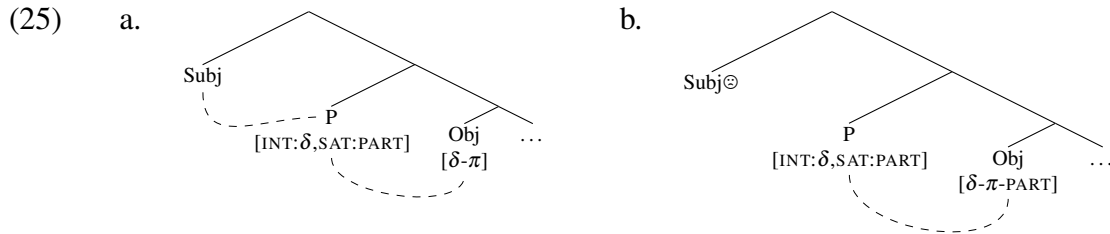
As noted in §2.1, the LDB in SSZ resembles the weak PCC, not the strong PCC: if there is one DP in a monotransitive, it must be in object position. Thus, in conjunction with the strong PCC operating over two clitics, the LDB additionally rules out *DP/PART and *DP/3 combinations. In Deal 2024, weak PCC effects are captured using the notion of *dynamic interaction*: a target of Agree may change the interaction conditions of the probe in the course of the derivation (see fn. 18).[18] For the LDB in SSZ, we may take the feature [π] to be dynamic (annotated as [π↑]). If the object bears [PART], then P is satisfied in

---

[18]Specifically, in the formulation of Deal 2024, the weak PCC can be captured if the probe P bears [INT:π,SAT:PART], but the [PART] feature is *dynamic* ([PART↑]): if P encounters a goal bearing [PART↑], then it gets updated to [INT:PART,SAT:PART]. This captures licit PART/PART combinations (e.g., 2/1, 1/2) but rules out illicit 3/PART combinations.

the first cycle of Agree (the strong PCC in (25b)). But if the object bears [$\pi\uparrow$], then P gets updated to [INT:π,SAT:PART]. This does not impact Agree with pronominal subjects (which already bear [π]), but does prevent Agree with lexical DP ([δ]) subjects, (26).
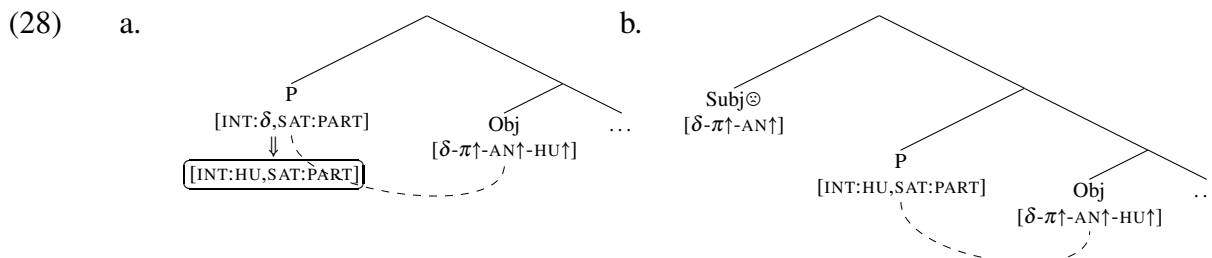
(26)
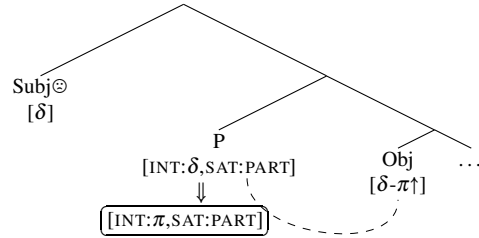
Finally, we turn to the GCC in SSZ: 3rd person clitics in SSZ display restrictions along a ranking of EL(der)>HU(man)>AN(imate)>IN(animate) (Foley and Toosarvandani, 2022; Foley et al., 2019), in that the first must outrank the second in animacy. This may be represented geometrically as additional specifications of [π], with higher-ranked features entailing lower-ranked ones (e.g., an elder pronoun bears [δ-π-AN-HU-EL], while an inanimate one bears only [δ-π]). As Foley and authors show, there is variation within SSZ as to exactly which clitic combinations induce the GCC; for space reasons, only the Yalálag variety is shown here, though see fn. 19 on other varieties. Representative examples are given in (27) (the full patterns can be found in Foley and Toosarvandani 2022, pp. 12-13).

(27) a. Bdin=**ba'**   **lebe'**   b. Wdao=**ba'**=**n**
CMPL.bite=3.AN 3.HU   CMPL.eat=3.AN=3.IN
'It (anim.) bit her/him.' (*3.HU =*be'*) 'It (anim.) ate it (inan.).'
(Foley and Toosarvandani, 2022, p. (López and Newberg, 2005, p. 8)
12)

In Sichel and Toosarvandani to appear, the GCC is derived with Probe Activation: a high probe P cannot undergo a second round of Agree with an object bearing animacy features that P has already copied from the subject. However, the GCC may also be derived in an int/sat model. Following Deal and Royer to appear on similar animacy-based restrictions in Mayan languages (see fn. 21 in §4.2), we may specify animacy features in SSZ, subfeatures

of [$\pi\uparrow$], as dynamic as well. Thus, if P interacts with an object with a dynamic animacy feature, e.g., [HU$\uparrow$], then P is updated to [INT:HU,SAT:PART], as in (28a). P cannot then Agree with a subject lacking this feature, (28b).[19]

(28) a.

P
[INT:δ,SAT:PART]
$\Downarrow$
[INT:HU,SAT:PART]

Obj
[δ-$\pi\uparrow$-AN$\uparrow$-HU$\uparrow$]
...

b.

Subj☺
[δ-$\pi\uparrow$-AN$\uparrow$]

P
[INT:HU,SAT:PART]

Obj
[δ-$\pi\uparrow$-AN$\uparrow$-HU$\uparrow$]
...

In contrast, SJPM lacks any dynamically interacting features. There are no restrictions on animacy (or noun class) combinations; for instance, human object clitics may co-occur with animate and inanimate subject clitics, (29).

(29) a. ko^1-ni^5-tu^{3}vi^3=**ti^5**=**ɲa^5**
        NEG-CMPL-sting=3.AN=3SG.F
        'It (the wasp) didn't sting her.'

    b. ka^{5}ʔno^3 ki^{5}ⁿdo^{1}o^3=**ɲa^1**=**ɲa^5**
        big   CONT.fit=3.N=3SG.F
        'It (the shirt) fits her big.'   (SJPM)

The int/sat model of Agree (Deal, 2024) thus captures the range of hierarchy effects in SJPM and SSZ in a mostly uniform manner. Both languages have a low probe P specified for [INT:δ,SAT:PART] that encounters the object before (possibly) Agreeing with the subject; the crucial distinction is that SSZ has dynamic [π] features (including animacy subfeatures) that may change P's interaction specification in the course of the derivation.

What happens *after* P fails to Agree with a subject? For pronominal subjects, the outcome is unsurprising: since clitics are generated by Agree, failure to Agree with the subject

---

[19]Beyond Yalálag, other SSZ varieties display slightly different versions of the GCC (Foley et al., 2019; Foley and Toosarvandani, 2022; Foley et al., to appear), though all varieties display the LDB. For instance, the Laxopa variety permits HU/EL clitic combinations in either order, though otherwise obeys the HU>AN>IN ranking of Yalálag; the Zoogocho variety is even more permissive: HU/EL and AN/{EL,HU} combinations are licit, though IN/AN is still ruled out. Under the present int/sat analysis, we can understand this variation as follows. Across SSZ, [$\pi\uparrow$] is invariably dynamic (accounting for the LDB), but there is variation in which animacy subfeatures of [π] are also dynamic. In Yalálag, all animacy features are dynamic; in Laxopa, [AN$\uparrow$] and [HU$\uparrow$] are dynamic but [EL] is not; and in Zoogocho, only [AN$\uparrow$] is dynamic. Notice that the choice of dynamic features is not random in these varieties. Given a common feature geometry of [$\pi\uparrow$-AN($\uparrow$)-HU($\uparrow$)-EL($\uparrow$)], the patterns discussed here instantiate three different points along an implicational hierarchy.

entails failure to generate a clitic. But it is less obvious why the inability to Agree with a lexical DP subject should yield ungrammaticality: as also pointed out in Sichel and Toosarvandani to appear, this does not automatically fall out from the system of Deal 2024, since lexical DPs are not generated by Agree and thus do not (all things being equal) *require* the probe P to Agree with them. This will be the topic of §5. However, we must first address the distribution of tonic pronouns, which connects to our treatment of lexical DPs.

4.2 *Tonic pronouns, lexical DPs, and the nominal spine*

If we maintain the assumption that Agree is needed for cliticization, Agree with only the object should prevent generation of a clitic in subject position. But only the opposite pattern is attested in both SJPM and SSZ: it is the object and not the subject that necessarily surfaces as a tonic pronoun, shown in (30) in SJPM.

(30) a. ko^5ni^1=$^n\textbf{do}^5$ $\boxed{^n\textbf{du}^1\textbf{(}^{\textbf{ʔ}}\textbf{u}^1\textbf{)}}$ b. *ko^5ni^1 $\boxed{^n\textbf{do}^5\textbf{(}^{\textbf{ʔ}}\textbf{o}^1\textbf{)}}$=$^n\textbf{di}^1$
  CONT.love=2PL 1PL.EX   CONT.love 2PL=1PL.EX
  'You (pl.) love us (excl.).'   Intended: 'You (pl.) love us (excl.).'

I propose that this problem is only apparent: pronominal clitics and tonic pronouns are not derivationally related. Rather, a pronominal clitic expones a π-containing feature bundle and may be represented as a (DP-less) ϕP (cf. Dechaine and Wiltschko, 2002), (31a), whereas a tonic pronoun is the realization of a ϕP enclosed within an outer DP shell, (31b). Thus, tonic pronouns are actually *portmanteaux*, spelling out the entire complex DP structure.[20] Crucially for our purposes, tonic pronouns are $[\delta]$ (like lexical DPs); their internal π-features are inaccessible to any Agreeing probes that may interact with them.

(31) a. *Pronominal clitic:* $\phi P_{[\delta-\pi-(...)]}$ b. *Tonic pronoun:* DP$_{[\delta]}$
                                D$_{[\delta]}$ $\phi P_{[\delta-\pi-(...)]}$

---

[20]This paper does not commit to a particular mechanism for portmanteau formation. For instance, a portmanteau may expone a non-terminal node (Caha, 2009) or consist of a span of contiguous heads (Merchant, 2015; Svenonius, 2016). It is also plausible that the proposed ϕP itself contains articulated functional structure encoding individual ϕ-features (e.g., person, animacy, number) in each terminal node.

There is independent evidence for the structure in (31b). First, for SSZ, recall that the GCC constrains certain combinations of 3rd person clitics. Importantly, the GCC only applies when *both* arguments are clitics: no GCC effects arise when one or both arguments is not a clitic. Recall from (27a) that an animate subject clitic may be outranked by a human tonic pronoun object; (32) shows the same with a human lexical DP object. Thus, the dynamic animacy features of lexical DPs and tonic pronouns must be invisible to the low probe P.[21]

(32) Nha'^{3} tso'o=**b^3**   **bi^{3}do'^{3} ka'^{1}**
      then CONT.PL.eat=3.AN child   those
      'Then they (animals) were eating those children.' (Laxopa; Toosarvandani 2023)

In addition, Sichel and Toosarvandani (2020) observe that all 3rd person pronouns in SSZ consist of a pronominal enclitic attached to a formative *le*, as summarized in (33) (see also Solá-Llonch 2021 for a similar conclusion for San Cristobal Lachirioag Zapotec). Thus, under the present approach, 1st/2nd person tonic pronouns are portmanteaux, while their 3rd person counterparts transparently show the decomposition shown in (31a).

(33)

	1SG	2SG	3.EL	3.HU	3.AN	3.IN
Clitic	=a'	=o'~=u'	=e'	=ba'	=(e)b~=ba	=(e)nh
Tonic pronoun	nada'~neda'	lhe'	**l=e'**	**le=ba'**	**le=b**	**l=enh**

(SSZ; adapted from Sichel and Toosarvandani 2020, p. 100)

The relevant evidence is difficult to detect in SJPM (there are no GCC effects, and there are no dedicated 3rd person tonic pronoun forms[22]). However, the proposed decomposition of

---

[21] We may moreover contrast the GCC in SSZ to similar animacy-based restrictions in Mayan languages, as recently analyzed by Deal and Royer (to appear). As (i) shows, combinations of lexical DPs are ruled out in Mayan languages such as Chuj (a VOS language) whenever the object outranks the subject (Deal and Royer also mention that pronominal objects may not outrank lexical DP subjects in animacy). Under the present account, the crucial difference amounts to whether the π-features remain within the ϕP (as in SSZ) or percolate up to the outermost DP layer (as in Mayan), such that they would be visible to the relevant probe.

(i)   a. Ix-y-il   **nok' chan winh winak**   b. *Ix-y-il   **winh winak nok' chan**
          PFV-A3-see CL snake CL   man   PFV-A3-see CL man   CL   snake
          'The man saw the snake.'   Intended: 'The snake saw the man.'
                                                        (Chuj; Deal and Royer to appear)

[22] In environments that disallow clitics (e.g., within coordinate structures and in non-neutral discourse

tonic pronouns is visible in other Mixtec varieties. As described in Mantenuto 2020, San Sebastián del Monte Mixtec (SSMM), very closely related to SJPM, also has the strong PCC and the LDB, the latter shown in (34). The object must be expressed with a tonic pronoun, just as in SJPM. Importantly, Mantenuto (2020) observes that 2SG.HON lacks a dedicated tonic pronoun portmanteau—but is still constrained by the strong PCC and LDB, (35). 2SG.HON objects are obligatorily expressed as pronominal clitics attached to the morpheme *mee*, whose cognates elsewhere in Mixtec have been treated as a definite marker (North, 1987; Cline, 2018; Hofmann and Ostrove, 2020).[23] Assuming that *mee* instantiates D, the proposed structure from (31b) is given in (36) for 2SG.HON:

(34) a. Kàni   Juan **yo'o**   b. Sàsi   Juan=**ti**
          CMPL.hit Juan 2SG   CMPL.eat Juan=3.AN
          'Juan hit you.' (*2SG =ó)   'Juan ate it.'

(35) Kàni   Juan **mee=ní**   (36)   DP$_{[\delta]}$
      CMPL.hit Juan BASE=2SG.HON   /\
      'Juan hit you (hon.).' (*2SG.HON =ní alone)   D$_{[\delta]}$ ϕP$_{[\delta-\pi-\text{PART}-\text{HON}]}$
      (Mantenuto, 2020, pp. 68, 70)   mee   ní

Thus, I conclude that pronominal clitics and tonic pronouns in SSZ and SJPM are syntactically distinct; either one may be Merged into the structure in a given derivation, modulo any grammatical (morphosyntactic or prosodic) constraints ruling out one or the other.[24] But whereas certain pronominal clitics in object position are later filtered out by the PCC and its kin, the probe P may interact with a $[\delta]$ tonic pronoun without being satisfied and may thus cyclically expand to probe for its specifier.

contexts), 3rd person pronominal meanings are encoded in SJPM using a prenominal classifier and a deictic, e.g., *ra¹ ka⁵a³* 'he (there)'. These clearly resemble the lexical DPs, which also co-occur with classifiers.

[23]Overall, there has been very little work on the syntacticosemantic properties of *mee* and its cognates; in most descriptive literature, this morpheme is (pretheoretically) referred to as a 'specifier'. This marker appears to be polyfunctional, in that it is also used to express emphasis and various other information structural meanings. Importantly, though, Mantenuto (2020) notes that its usage in (35) does not yield an emphatic reading—since this is the only way to express a 2SG.HON pronoun in object position.

[24]Indeed, in the absence of constraints such as the PCC and LDB, SJPM permits both clitic and tonic pronouns in neutral discourse contexts (the tonic pronouns usually surface in their weak or reduced forms, as the non-reduced forms are associated with focus or emphasis, as in (23)).Both are routinely accepted by speakers in formal elicitation settings as optional variants, and are often interchangeable in narratives.

5 The LDB as a subject licensing failure

I have proposed that violations of LDB result from failure of a low probe P to undergo second cycle Agree with a lexical DP subject, P's specifier. This section develops this further: although P may fail to Agree with the lexical DP (see also Preminger 2011, 2014), what is ultimately yields ungrammaticality is the fact that the lexical DP does not interact with *any* Agreeing head in the course of the derivation. The LDB is thus an interaction between the PCC and the Case Filter, the abstract requirement that all DPs are case-licensed by some A-probe via Agree (Vergnaud 1977; Chomsky 2000). Crucially, there is no additional probe P' (represented in (37) with [INT:δ:SAT:δ]) that may license a subject in P's stead.

(37)   *Unattested:*
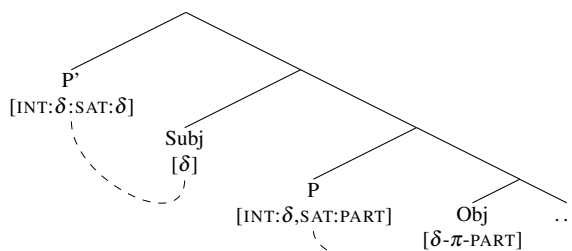
Crucially, the non-existence of additional subject licensors is directly connected to the verb-initial structure of SJPM and SSZ, which places all licensing heads in the VoiceP domain. In §5.1, I just focus on SJPM, but the points below may hold for SSZ as well, based on existing work on Zapotec syntax (e.g. Lee, 2006; Sonnenschein, 2004); §5.2 then discusses the predictions of this account for the LDB beyond Eastern Otomanguean.

5.1   V(P)-S-O word order and the loci of A-probes

It was noted in §2.1 that deriving V(P)-S-O word order involves object shift and remnant VP-fronting to a pre-subject position (cf. Massam, 2001). If we assume that the object raises to Spec-*v*P and the subject is generated in a higher Spec-VoiceP (Pylkkänen, 2002; Harley, 2013), we may update the schema in (5) as (38). Thus, our low probe P responsible for the PCC is in Voice. I also assume that the head F, which triggers VP-movement to its

specifier in (38), is a distinct head beneath T (van Urk, 2024): tense is realized at the left edge of the raised verb (whether as grammatical tone or as a segmental morph).[25]

(38)
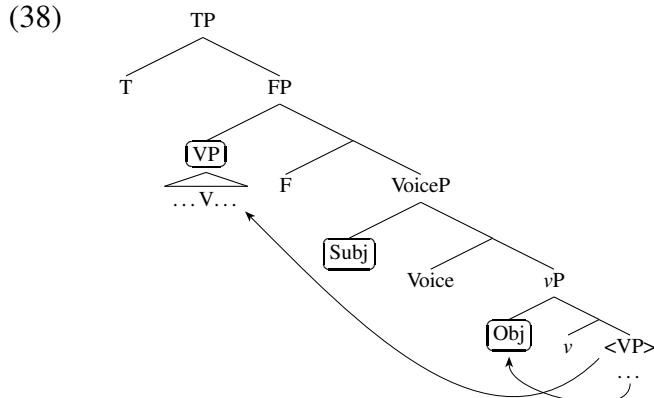
Putting together (37) and (38), the emergent system is that the Case Filter must be satisfied within the VoiceP domain. Voice and *v* may license DPs upon Agree, while higher heads such as F and T (and possibly others not represented in (38)) do not. This asymmetry is not incidental: much prior work on verb-initial languages has sought to parameterize the probe specification of T (or F in SJPM) so that it seeks a verbal element rather than a DP (Alexiadou and Anagnostopoulou, 1998; Oda, 2005; Potsdam, 2009; Doner, 2019). In SJPM, F seeks verbal and non-verbal (e.g., adjectival) predicates alike, but not DPs. Likewise, we will see that, while T hosts tense distinctions,[26] T does not serve additional grammatical functions associated with DP-licensing. To see this, let us (maximally generously) assume that a possible DP-licensor could be any functional head bearing an A-probe (e.g., bearing ϕ or case features or triggering A-movement). As SJPM systematically lacks overt case on DPs and ϕ-agreement morphology (beyond the pronominal clitics), we may detect the occurrence of licensing with A-movement and, of course, the PCC.

---

[25]Although the basic logic of this section is also compatible with a head movement derivation, as proposed for other Mesoamerican languages (Macaulay, 2005; Clemens and Coon, 2018), in SJPM and closely related varieties there is much evidence that the fronted verbal element is phrasal (e.g. Hedding and Yuan, to appear).

[26]For the purposes of this discussion, TAM distinctions (labeled as completive, continuative, and irrealis respectively, using Mixtecan glossing conventions) are framed as tense (Hollenbach, 2015), though they are sometimes characterized as aspect (e.g. Bickford and Marlett, 1988). Under an aspect-based treatment, the point remains: there is no correlation between TAM and the distribution of DPs.

First, SJPM seems to lack derived A-positions external to VoiceP. For instance, it wholly lacks expletive subjects and constructions that could be analyzed as raising-to-subject.[27] As exemplified in (39), SJPM speakers routinely provide paraphrases in direct translation tasks, suggesting the non-existence of such constructions. Likewise, it is not possible to raise a DP to a position immediately below C, (40), suggesting that T lacks an EPP feature.

(39) tu^{5}va^{1}ʔa^{3}ɾa^3 ni^1-kõ^{3}õ^3 sa^{1}vi^5
probably   CMPL-fall rain
'Probably rain fell.' (Targeted: 'Parece que llovió' or 'It seems to have rained.')

(40) *a^5 **ɾa^1**   **ⁿdʒaʃ^{51}**$_i$ ʃa^{13}ʔa^3 t_i ĩ^3 ʒu^{5}ʃi^{1}ni^3 ka^{5}ʔno^3 ⁿda^5 ɲa^5
Q CL.3SG.M Josh   CMPL.give   one hat   big   to CL.3SG.F
Clau^{5}dia^6?
Claudia
Intended: 'Did Josh give a big hat to Claudia?'
[Second author]'s comment about judgment: 'Ah, no... primero "a^5 ʃa^{13}ʔa^3."'

Although short (often string-vacuous) A-movement of subject DPs to a position below the raised verb has been posited for other verb-initial languages (e.g. Longenbaugh and Polinsky, 2018; Adler et al., 2018),[28] there is no positive evidence for this in SJPM: for instance, there are no adverbs or overt heads that may intervene between the subject and object. And although DPs may occur preverbally in SJPM (in the absence of overt C), this involves Ā-positions, generally associated with quantifier raising or information structure (Macaulay, 2005). Indeed, we may contrast Ā-movement with object shift, i.e., genuine A-movement. In SJPM, object shift is evidenced by a pattern of optional quantifier float, (41a), whereby a quantifier associated with the object may remain within the VP (which is then fronted to Spec-FP). Importantly, the quantifier cannot be stranded by a preverbal

---

[27]Moreover, as shown in Ostrove 2021 on the closely related San Martín Peras variety of Mixtec, passive constructions are impersonals, so they do not involve A-movement of the theme to a Voice-external position.

[28]In their proposed analysis of SSZ clause structure, Adler et al. (2018) do posit a short A-movement step of subjects outside of the extended *v*P domain, due to the relative word order of nominal arguments and *v*P-internal manner adverbs. However, the data provided are also compatible with the present account, which diverges from Adler et al. 2018 by decomposing the *v*P into distinct projections, VoiceP (the locus of the subject) and *v*P (within which manner adverbs are adjoined).

subject, (41b). For many other languages, quantifier float is taken to be licensed by A-movement but not Ā-movement (Sportiche, 1988; Zyman, 2018). Extending this idea to SJPM, this reifies that the preverbal subject position in (41b) is not an A-position.

(41) a. [$_{VP}$ ʃa^{13}ʃi^5 {^{n}di$^{3?}$i^3}]=ɲa^5 [{^{n}di$^{3?}$i^3} **ɲa^1** **ti^{1}ko^{3}o^{13}**]
     CMPL.eat all =3SG.F all CL.3.N tamale
     'She ate all of the tamales.'

   b. [**^{n}di$^{3?}$i^3 ʒo^{5}o^5**] ʃi^{5}tʃi^3
      all 1PL.IN CONT.swim/bathe
      'All of us are swimming/bathing.' (Not possible: *ʒo^{5}o^5_i ʃi^{5}tʃi^3 [^{n}di$^{3?}$i^3 t_i])

Finally, in SJPM, we do not find familiar hallmarks of DP-licensing by finite T. SJPM lacks dedicated non-finite morphology, but the so-called 'irrealis' verbal form is used in obligatory control and other contexts that are naturally translated as non-finite (Ostrove, to appear). In such contexts, the embedded subject is necessarily overt.[29] More generally, there are (to my knowledge) no constructions in the language that restrict the distribution of subjects based on TAM distinctions.

(42) a. ku^{5}u^3 ɲa^5 si$^{5?}$i^6 [ka^{3}ta^3=**ɲa^5**]
     CONT.try CL.3SG.F woman IRR.sing=3SG.F
     'The woman is trying to sing.'

   b. ko^{13}ni^1=ɲa^5 [ka^{3}ʃa$^{5?}$a^3 **ɾa^1** **X^{w}ã^{5}ã^1**]
      CMPL.want=3SG.F IRR.eat CL.3SG.M Juan
      'She wanted Juan to eat.'

Altogether, these considerations provide converging evidence for the absence of any VoiceP-external A-probes. An object may be licensed by *v*, raising it to Spec-*v*P (where it may then interact with Voice). However, a subject in Spec-VoiceP may only be targeted by Voice—provided that Voice may undergo second-cycle Agree.

---

[29]Ostrove (to appear) analyzes the embedded pronominal subject in (ia) as overt PRO (specifically, pronominal copy control) in closely related San Martín Peras Mixtec.

5.2 *The (absence of) the LDB cross-linguistically*

While other languages display PCC effects among monotransitive arguments, the LDB is (thus far) only found in SSZ and SJPM. We may wonder, however, why it is not attested more widely. I suggest that this is due to its *conspiratorial* nature, arising only when highly specific conditions happen to align in a given language: (i) the Case Filter, (ii) a low probe P with a [SAT:PART] condition, and (iii) no other A-probes available to license a DP that P cannot interact with. Other languages may simply lack properties (ii) and/or (iii).

That being said, as noted in fn. 13, a identical to the LDB in SJPM does hold outside of monotransitives in certain Romance languages, a phenomenon known as the 'Fancy Constraint' (Postal, 1989; Rezac, 2011; Sheehan, 2020; Deal, 2024): in *faire-infinitif* causatives containing a dative causee, which may be a clitic (not shown) or a lexical DP, the direct object may not be a [PART] clitic, (43). This is distinct from the canonical PCC pattern that holds of ditransitive internal arguments, which does not affect lexical DPs.

(43) Marcel **l'**   / ***vous**   a   fait   épouser **au** **médecin**
      Marcel 3.ACC= 2PL.ACC= has made marry to.the doctor
      'Marcel had the doctor marry her / *you.'   (French; Sheehan 2020, pp. 151-152)

Here, I offer a potential new perspective on the Fancy Constraint (to be developed in future work). Following Folli and Harley 2007, the causee of a *faire-infinitif* is introduced as the specifier of *v*P, directly embedded under the causative verb *faire* (which introduces the causer).[30] Extending our analysis of SJPM, a causee in Spec-*v*P may fail to be licensed if *v* bears [SAT:PART] and is satisfied in its first cycle of Agree. Unlike in SJPM, however, higher licensing heads such as T do exist in French—but T invariably Agrees with the external argument, which is more local. Thus, Romance *faire-infinitifs* may be another syntactic configuration in which the aforementioned properties (i)-(iii) simultaneously hold.

---

[30]I also assume, following Deal 2024, that DAT morphology in (43) simply signals successful Agree between *v* and its specifier.

6 Conclusion

This paper has presented an alternative to Sichel and Toosarvandani's (to appear) analysis of the LDB in SSZ, which takes lexical DP subjects to block a high probe from licensing object clitics. I have argued for the opposite logic, based on new comparative evidence from SJPM: objects may prevent a low probe from licensing lexical DP subjects.

I have also argued that the variation between SSZ and SJPM should be analyzed as micro-parameterization within an otherwise common grammatical system. In this paper, the crucial difference lies simply in the inventory of dynamically interacting π-features in SSZ but not in SJPM. Alternatively, we may model the variation by appealing to more significant differences, for instance by varying the height of the probe (high in SSZ but low in SJPM) or the featural content of lexical DPs ($[\delta]$ in SSZ but $[\pi]$ in SJPM?). However, maintaining a near-uniform syntax has the benefit of narrowing the hypothesis space and making stronger predictions that may be evaluated in future work.

References

Adler, Jeff, Steven Foley, Jed Pizarro-Guevara, Kelsey Sasaki, and Maziar Toosarvandani. 2018. The derivation of verb initiality in Santiago Laxopa Zapotec. In *A reasonable way to proceed: Essays in honor of Jim McCloskey*, ed. Jason Merchant, Line Mikkelsen, Deniz Rudin, and Kelsey Sasaki, 31–49. Santa Cruz, Berkeley & Chicago.

Alexander, Ruth. 1988. A syntactic sketch of Ocotepec Mixtec. In *Studies in the syntax of mixtecan languages*, ed. C. Henry Bradley and Barbara E. Hollenbach, 151–304. Dallas, TX and Arlington, TX: SIL and the University of Texas at Arlington.

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16:491–539.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Berlin: Mouton de Gruyter.

Avelino Becerra, Heriberto. 2004. Topics in Yalálag Zapotec, with particular reference to its phonetic structure. Doctoral Dissertation, UCLA, Los Angeles, CA.

Baker, Mark, and Ruth Kramer. 2018. Doubled clitics are pronouns: Amharic objects and beyond. *Natural Language and Linguistic Theory* 36:1035–1088.

Béjar, Susana. 2003. Phi-syntax: A theory of agreement. Doctoral Dissertation, University of Toronto.

Béjar, Susana, and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: Theory and acquisition*, ed. Ana-Teresa Perez-Leroux and Yves Roberge, 49–61. Amsterdam: John Benjamins.

Béjar, Susana, and Milan Rezac. 2009. Cyclic Agree. *Linguistic Inquiry* 40:35–73.

Bickford, Albert J., and Stephen A. Marlett. 1988. The semantics and morphology of Mixtec mood and aspect. *Work Papers of the Summer Institute of Linguistics* 32:1–39.

Bonet, Eulalia. 1991. Morphology after syntax: Pronominal clitics in Romance. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Caballero, Gabriela, Claudia Duarte Borquez, Claudia Juárez Chávez, and Michelle Yuan. to appear. Lexical and grammatical tone in San Juan Piñas Mixtec (Tò'ōn Ndá'ví). *Phonological Data and Analysis* .

Caha, Pavel. 2009. The nanosyntax of case. Doctoral Dissertation, University of Tromsø, Tromsø.

Campbell, Eric. 2017a. Otomanguean historical linguistics: Exploring the subgroups. *Language and Linguistics Compass* 11:e12244.

Campbell, Eric. 2017b. Otomanguean historical linguistics: Past, present, and prospects for the future. *Language and Linguistics Compass* 11:e12240.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the languages of Europe*, ed. Henk van Riemsdijk, 145–233. New York: Mouton de Gruyter.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Clemens, Lauren Eby, and Jessica Coon. 2018. Deriving verb-initial word order in Mayan. *Language* 94:237–280.

Cline, Laura. 2018. Some features of participant reference in Xochapa Mixtec. Master's thesis, University of North Dakota, Grand Forks, ND.

Coon, Jessica, and Stefan Keine. 2021. Feature gluttony. *Linguistic Inquiry* 52:665–710.

Deal, Amy Rose. 2015. Interaction and satisfaction in ϕ-agreement. In *Proceedings of NELS 45*, ed. Thuy Bui and Deniz Ozyildiz, 179–192. Amherst, MA: GLSA.

Deal, Amy Rose. 2024. Interaction, Satisfaction, and the PCC. *Linguistic Inquiry* 55:39–94.

Deal, Amy Rose, and Justin Royer. to appear. Mayan animacy hiearchy effects and the dynamics of Agree. *Natural Language and Linguistic Theory* .

Dechaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33:409–442.

Doner, Julianne. 2019. The EPP across languages. Doctoral Dissertation, University of Toronto, Toronto, ON.

Duarte Borquez, Claudia. 2022. The representation of tone in San Juan Piñas Mixtec. Comprehensive paper, UC San Diego.

Foley, Steven, Nick Kalivoda, and Maziar Toosarvandani. 2019. Forbidden clitic clusters in Zapotec. In *Proceedings of CLS 53*, ed. Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit, and Michael Tabatowski, 87–102. Chicago: Chicago Linguistics Society.

Foley, Steven, Nick Kalivoda, and Maziar Toosarvandani. to appear. Gender-Case Constraints in Zapotec. In *Proceedings of WSCLA 22*, ed. Anne Bertrand. Vancouver, BC: UBCWPL.

Foley, Steven, and Maziar Toosarvandani. 2022. Extending the Person–Case Constraint to gender: Agreement, locality, and the syntax of pronouns. *Linguistic Inquiry* 53:1–40.

Folli, Raffaella, and Heidi Harley. 2007. Causation, obligation, and argument structure: On the nature of little v. *Linguistic Inquiry* 38:197–238.

Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory* 32:1033–1088.

Harley, Heidi. 2013. Getting morphemes in order: Merger, affixation, and head movement. In *Diagnosing syntax*, ed. Lisa Cheng and Norbert Hornstein, 44–74. Oxford: Oxford University Press.

Harley, Heidi, and Elizabeth Ritter. 2002. A feature-geometric analysis of person and number. *Language* 78:482–526.

Hedding, Andrew A., and Michelle Yuan. to appear. Phase unlocking and the derivation of verb-initiality in San Martín Peras Mixtec. In *Proceedings of WCCFL 41*. Somerville, MA: Cascadilla Press.

Hofmann, Lisa, and Jason Ostrove. 2020. Míí and classifiers in smpm. Presentation at WLMA, UC Santa Cruz.

Hollenbach, Barbara E. 2015. Notes on tense, mood, and negation in Mixtec languages. Ms., Summer Insititute of Linguistics (SIL) International.

Josserand, Judy Kathyrn. 1983. Mixtec dialect history. Doctoral Dissertation, University of Michigan Ann Arbor, Ann Arbor, MI.

Kaufman, Terrence. 1988. Otomanguean tense, aspect, mood, voice, and nominalization markers. Ms., University of Pittsburgh.

Lee, Felicia. 2006. *Remnant raising and VSO clausal architecture*. Dordrecht: Springer.

Longenbaugh, Nicholas, and Maria Polinsky. 2018. Equidistance returns. *The Linguistic Review* 35:413–461.

López, Filemón, and Ronaldo Newberg. 2005. *La conjugación del verbo zapoteco, zapoteco de Yalálag*. Mexico City: Instituto Lingüístico de Verano.

Lopez, Oswaldo Cruz. 2022. Cláusulas de relativo en el Tu'un Savi de Santo Tomás Ocotepec, tlaxiaco, oaxaca. Master's thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, Mexico City.

Macaulay, Monica. 1987. Cliticization and the morphosyntax of Mixtec. *International Journal of American Linguistics* 53:119–135.

Macaulay, Monica. 1996. *A grammar of Chalcatongo Mixtec*. Berkeley, CA: University of California Press.

Macaulay, Monica. 2005. The syntax of chalcatongo mixtec: Preverbal and postverbal. In *Verb first: On the syntax of verb-initial languages*, ed. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 341–366. Amsterdam: John Benjamins.

Mantenuto, Iara. 2020. 'other' and 'more' in San Sebastián del Monte Mixtec and beyond. Doctoral Dissertation, UCLA, Los Angeles, CA.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19:153–197.

Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46:273–303.

Nevins, Andrew. 2007. The representation of third person and its consequences for Person-Case effects. *Natural Language and Linguistic Theory* 25:273–313.

Nevins, Andrew. 2011. Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language and Linguistic Theory* 29:939–971.

North, Joanne. 1987. Participant tracing and prominence in Silacayoapan Mixtec. *SIL Mexico Workpapers* 8:100–128.

Oda, Kenji. 2005. V1 and wh-questions: A typology. In *Verb first: On the syntax of verb-initial languages*, ed. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 107–134. Amsterdam: John Benjamins.

Ostrove, Jason. 2021. Impersonals in San Martín Peras Mixtec. In *Proceedings of WSCLA 25*, ed. Marianne Huijsmans, D. K. E. Reisinger, and Rose Underhill, 116–130. Vancouver, BC: UBCWPL.

Ostrove, Jason. to appear. Obligatorily overt PRO in San Martín Peras Mixtec. *Linguistic Inquiry* .

Pancheva, Roumyana, and Maria Luisa Zubizarreta. 2018. The Person Case Constraint: The syntactic encoding of perspective. *Natural Language and Linguistic Theory* 36:1291–1337.

Penner, Kevin L. 2019. Prosodic structure in Ixtayutla Mixtec: Evidence for the foot. Doctoral Dissertation, University of Alberta, Edmonton, AB.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Postal, Paul. 1989. *Masked inversion in French*. Chicago: University of Chicago Press.

Potsdam, Eric. 2009. Malagasy backward object control. *Language* 85:754–784.

Preminger, Omer. 2011. Agreement as a fallible operation. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.

Pylkkänen, Liina. 2002. Introducing arguments. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Rezac, Milan. 2011. *Phi-features and the modular architecture of language*. Dordrecht: Springer.

Sheehan, Michelle. 2020. The Romance person case constraint is not about clitic clusters. In *Dative constructions in Romance and beyond*, ed. Anna Pineda and Jaume Mateau, 143–171. Berlin: Language Science Press.

Sichel, Ivy, and Maziar Toosarvandani. 2020. Pronouns and attraction in Sierra Zapotec. *Syntax at Santa Cruz* 4:97–116.

Sichel, Ivy, and Maziar Toosarvandani. to appear. The featural life of nominals. *Linguistic Inquiry* .

Sichel, Ivy, and Martina Wiltschko. 2021. The logic of Person markedness: Evidence from pronominal competition. *Language* 97:42–71.

Solá-Llonch, Elizabeth. 2021. Aspects and arguments of the Lachirioag Zapotec verb. Master's thesis, UCLA, Los Angeles, CA.

Sonnenschein, Aaron Huey. 2004. A descriptive grammar of San Bartolomé Zoogocho Zapotec. Doctoral Dissertation, University of Southern California, Los Angeles, CA.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425–449.

Sportiche, Dominique. 1993. Clitic constructions. Ms., University of California Los Angeles.

Stegovec, Adrian. 2020. Taking case out of the Person-Case Constraint. *Natural Language and Linguistic Theory* 38:261–311.

Svenonius, Peter. 2016. Spans and words. In *Morphological metatheory*, ed. Heidi Harley and Daniel Siddiqi, 199–220. Amsterdam: John Benjamins.

Toosarvandani, Maziar. 2017. On reaching agreement early (and late). In *Asking the right questions: Essays in honor of Sandra Chung*, ed. Jason Ostrove, Ruth Kramer, and Joseph Sabbagh, 124–138. Santa Cruz, CA: UCSC Linguistics.

Toosarvandani, Maziar. 2023. The interpretation and grammatical representation of animacy. *Language* 99:760–808.

van Urk, Coppe. 2024. Constraining VP-fronting. *Linguistic Inquiry* 55:327–373.

Vergnaud, Jean-Roger. 1977. Personal letter to H. Lasnik and N. Chomsky.

Zylstra, Carol F. 2012. *Gramática del Tu'un Savi (la lengua mixteca) de Alacatlatzala, Guerrero*. Tlalpan, D.F., México: Instituto Lingüístico de Verano.

Zyman, Erik. 2018. Quantifier float as stranding: Evidence from Janitzio P'urhepecha. *Natural Language and Linguistic Theory* 36:991–1034.