

# The French demonstrative paradigm: Structurally transparent but semantically intricate

Adèle Hénot-Mortier

Massachusetts Institute of Technology

## 1. Introduction

French demonstratives form a comprehensive paradigm, already described in Bernstein (1997) and Kayne and Pollock (2010). This paradigm involves a demonstrative determiner (henceforth, **DEM**) realized as *ce*, *cet* or *cette*<sup>1</sup> and two suffix-like markers *-ci* and *-là*. These markers, which are probably derived from the adverbs *ici* ('here') and *là* ('there'), can be used to express the proximal/distal distinction (often together with pointing), contrastive focus, or anaphoricity. We will gloss them as respectively **HERE** and **THERE**, even when they are intended to express something else than the proximal/distal distinction. The next couple of sections illustrate how the DEM+(-*ci*/-*là*) complex can combine with various kinds of syntactic objects, which in turn leads to intricate constraints on the availability of further modification by, e.g., relative clauses and CPs.

### 1.1 Adnominal demonstratives

(1a) shows how DEM can combine with a nominal. *-ci* and *-là* are both optional in that case and, when used, lead to a proximal/distal contrast. The additional relative clause (henceforth abbreviated **RC**) is optional, as well, and, when used along with *-ci* or *-là*, is necessarily non-restrictive (Bernstein 1997). In (1b) the occurrences of DEM combine with two different nominals. *-ci* and *-là* are again both optional, but, if present, are preferably used in that order, and have the effect of reinforcing contrastive focus, without necessarily indicating a proximal/distal distinction.

---

<sup>1</sup>The variation between those three written forms depend on the grammatical gender of the target of the determiner (typically, a noun) and the phonology of the next word (which in certain cases can be an adjective, for instance). DEM is realized as *ce* (/sə/) when combined with a masculine noun, and when the word immediately following DEM starts with a consonant. DEM is realized as *cet* (/set/) when combined with a masculine noun, and when the word immediately following DEM starts with a vowel. Lastly, DEM is realized as *cette* (/set/) when combined with a feminine noun, regardless on the phonology of the next word.

- (1) a. Marie aime **ce gars** ({-ci, -là}) (qui \_ lit un livre)<sub>RC</sub>.  
 Marie likes DEM guy ({-HERE, -THERE}) (who \_ reads a book)<sub>RC</sub>.  
 ‘Marie likes this/that guy who reads a book.’  
 b. Marie aime **ce gars (-ci)** mais pas **cette fille(-là)**.  
 Marie likes DEM guy (-HERE) but not DEM girl(-THERE).  
 ‘Marie likes this guy, but not this girl.’

## 1.2 Adpronominal demonstratives (DEM+strong pronoun)

DEM can also combine with third person strong pronouns in the sense of Cardinaletti and Starke (1999). Throughout this paper, strong forms will be glossed as STR, and weak forms as WK. In (2), the third person singular masculine pronoun *lui*, which is unambiguously strong (unlike for instance its feminine counterpart *elle*), combines with DEM. In that case, *-ci/-là*, some kind of complement of the demonstrative (e.g. *de son frère* in (2c)), or a RC *must* be realized. The RC and either *-ci/-là* or a complement of the demonstrative *can* be realized at the same time, leading to a non-restrictive reading of the RC. However, *-ci/-là* and the complement of the demonstrative *cannot* be both realized, which probably means that these two kinds of syntactic objects compete with each other in terms of argument structure. Note that (2a) is fine but DEM+pro<sub>strong</sub> forms are overall preferable with inanimate referents, as in (2b). We will come back to this question in Section 3.

- (2) a. Marie aime **celui** ({-ci, -là}) (qui \_ lit un livre)<sub>RC</sub>.  
 Marie likes DEM-3.SG.STR ({-HERE, -THERE}) (who \_ reads a book)<sub>RC</sub>.  
 ‘Marie likes this/that one who reads a book.’  
*Note: the sentence sounds a bit derogatory/objectifying towards the book reader.*  
 b. Marie aime **celui** ({-ci, -là}) (qui \_ est sucré)<sub>RC</sub>.  
 Marie likes -3.SG.STR ({-HERE, -THERE}) (which \_ is sweet)<sub>RC</sub>.  
 ‘Marie likes this/that one that is sweet.’  
*Note: the sentence sounds fine since celui can easily be considered an object (e.g. a pastry, a dessert...) in that context.*  
 c. Marie a un sac mais préfère **celui** de son frère.  
 Marie has a bag but prefers DEM-3.SG.STR of her brother.  
 ‘Marie has a bag but prefer that of her brother.’

The two tables below spell out all the possible combination of DEM with strong pronouns (3), and how they combine with *-ci/-là*, complements, and RCs (4).

(3) The DEM+pro<sub>strong</sub> paradigm.

	Singular	Plural
Feminine	<i>ce+elle=celle</i>	<i>ce+elles=celles</i>
Masculine	<i>ce+lui=celui</i>	<i>ce+eux=ceux</i>

(4) Possible combinations of markers and modifiers with DEM+pro<sub>strong</sub>.

<i>-ci/-là</i>	complement	RC
✓	✗	✓/✗
✗	✓	✓/✗
✗	✗	✓

### 1.3 “Bare” demonstratives

A more intricate picture emerges when DEM does not combine with any (overt) argument, as shown in (5-6). When *-ci* either or *-là* gets used as in (5a) and (5b), the referent is preferably eventive/propositional (in other words, abstract) and *can* be followed by a CP (but not a RC). When neither *-ci* nor *-là* is realized as in (6a) or (6b), the referent can either be an abstract or concrete individual, and *must* be followed by either a RC or a CP.

- (5) a. Marie aime **ce**{**ci**, **là**} (\*que Jean lit \_\_)<sub>RC</sub>  
Marie likes DEM{HERE, THERE} (\*that Jean reads \_\_)<sub>RC</sub>  
Intended: ‘Marie like this/that thing that Jean is reading.’  
b. Marie aspire à **ce**{**ci**, **là**} (que Jean lise)<sub>CP</sub>  
Marie strives for DEM{HERE, THERE} (that Jean read.SBJV)<sub>CP</sub>  
‘Marie strives for Jean to read.’
- (6) a. Marie aime **ce** \*(que Jean lit \_\_)<sub>RC</sub>  
Marie likes DEM \*(that Jean reads \_\_)<sub>RC</sub>  
‘Marie likes the thing that Jean reads.’  
b. Marie aspire à **ce** \*(que Jean lise)<sub>CP</sub>  
Marie strives for DEM \*(that Jean read.SBJV)<sub>CP</sub>  
‘Marie strives for Jean to read.’

In the next section, we show how most of the structures presented in the above can be transparently accounted for within the framework of Ahn (2022), bringing empirical support for this theory of demonstratives. In section 3, we will discuss secondary puzzles at the syntax-semantics interface posed by the DEM+pro<sub>strong</sub> and the “bare DEM” cases.

## 2. Account

### 2.1 Background

Ahn (2022) develops a unified theory of demonstratives making use of a binary maximality operator (**bi-sup**) taking two arguments: a set of restrictions, and a relation (**R**), which according to Ahn can be either one of a deictic pointing, an anaphoric index, or a RC. The Logical Form (LF) of demonstratives is illustrated in (7), taken from Ahn (2022). As shown in (7a), the entry for bi-sup is a function that takes two predicates (restrictions and relation) and returns the supremum of the individuals verifying both predicates. As shown in (7b) and (7c), the two predicates can have a wide range of denotations, depending on their syntactic realization.

If the restrictions slot is occupied by a *null* pronoun, the corresponding entry will simply denote an entity (7b-i). If instead, it is realized as an *overt* pronoun, the corresponding entry will denote the set of individuals possessing the  $\Phi$ -features expressed by the pronoun (7b-

ii). Lastly, if the restrictions slot hosts an NP, its denotation will simply be that of the NP (7b-iii). Turning to the relation predicate, it can denote a set of individuals located at the place of pointing (7c-i), a set of individuals equal to the reference of a specific index (7c-ii), or a predicate corresponding to the denotation of a RC (7c-iii).

(7) *The LF of demonstratives (Ahn 2022):*

$$\begin{array}{c}
 \text{DP} \\
 \swarrow \quad \searrow \\
 \text{D}' \quad \text{R} \\
 \swarrow \quad \searrow \\
 \text{bi-sup} \quad \text{restrictions}
 \end{array}$$

a.  $\llbracket \text{bi-sup} \rrbracket = \lambda P. \lambda R. \iota x. \forall y. P(y) \wedge R(y) \iff y \sqsubseteq x$

b.  $\llbracket \text{restrictions} \rrbracket = \begin{cases} \lambda x. \text{entity}(x) & (i) \\ \text{or } \lambda x. \text{feminine}(x) \wedge \text{singular}(x) & (ii) \\ \text{or } \lambda x. \text{student}(x) & (iii) \\ \text{or } \dots \end{cases}$

c.  $\llbracket \text{R} \rrbracket = \begin{cases} \rightarrow_L \triangleq \lambda x. x \text{ is at location } L^2 & (i) \\ \text{or } \text{idx}_i \triangleq \lambda x. x = g(i) & (ii) \\ \text{or } \lambda x. x \text{ reads a book} & (iii) \\ \text{or } \dots \end{cases}$

This unified account is not directly supported by a language like English, in which demonstratives such as *this* and *that* cannot be clearly decomposed into a bi-sup and a R-morpheme.<sup>3</sup> We will see in the next section that French, *via* its demonstrative paradigm, transparently instantiates the relevant structures and functions.

## 2.2 Proposal

We propose that, in the case of French, DEM (*ce*) and *-ci/-là* respectively fill the bi-sup and R slot of the LF in (7).

(8) *Mapping between ce, ci, and là and the LF (7):*

- a. **bi-sup** = DEM = *ce*
- b. **R**  $\in \{ci, là, RC^4\}$

More specifically, we assume that depending on the context, *-ci* and *-là* behave as linguistic reflexes of pointing, or introduce bound variables. We think this is possible, because at a certain level of abstraction those operations are the same: they equate an individual with something located at a certain place in the actual world, or within an abstract register, namely the assignment function *g*. We therefore define the entries for *-ci* and *-là* as abstract,

<sup>2</sup>  $\triangleq$  here means “as per Ahn’s definition”.

<sup>3</sup> One could argue that *(over) here* and *(over) there* are instances of R in English. However, those expressions cannot express the full range of meanings proposed for R. More specifically, they can express neither pure contrastive focus (without a proximal/distal distinction), nor anaphoric indexing.

<sup>4</sup> We treat complements of the demonstrative (such as *de son frère* in (2c)) as RCs filling the R-slot. This is justified by the fact that those complements compete with the markers *-ci* and *-là* and restrictive RCs, and, type-wise, can be seen as predicates just like RCs.

- Examples (10-12) illustrate the claims that *-ci* and *-là* realize the functions mentioned in the above. In (10), which is set in a context involving pointing, *-ci/-là* are mandatory with the DEM+pro<sub>strong</sub> construction.<sup>5</sup> The referent of *celle-cel/elle-là* must match the location denoted by the pointing gesture in a one-to-one fashion, consistent with the meaning of the *idx<sub>i</sub>* function. Examples (a-d) show that the proximal/distal contrast between *-ci* and *-là* is not absolute (in the sense that the distal marker can be used in all cases, cf. (10c)); but instead, relative: whenever the *-ci* marker is used, *-là* should be used as well (cf. 10d), and all locations signaled with *-ci* should be closer to all locations signaled with *-là*, cf. (10e). Example (10f) moreover confirms that the badness of (10e) is not solely due to having the *-là* marker linearly precede the *-ci*-markers.<sup>6</sup>

- <sup>5</sup>Note that this would not have been the case if *celle* had been replaced by a *ce*-NP construction.

<sup>6</sup>We associate the relative oddness of (10f) to external factors such as salience. More specifically, a speaker may preferably point to the most salient elements first, and salience happens to correlate with proximity.

(11) emphasizes the use of the *-ci/-là* alternation to mark contrastive focus. The markers appear again obligatory; pointing can target abstract *loci*. (12) shows how *-ci/-là* allows to track referents via binding, reminiscent of the English construction *the former ... the latter*.

- (11) *Context: Marie is again at a bakery. This time she inquires about éclairs. The bakery employee shows her that éclairs come in two sizes, big or small. The big and small éclairs are roughly at the same location.*

Celui-ci<sub>→1</sub> est GRAND, alors que celui-là<sub>→2</sub> est PETIT.  
 DEM-him-HERE is BIG, while COMP DEM-him-THERE is SMALL.  
 ‘This one is big, while this other one is small.’

- (12) Chaque fois que je vois un chien<sub>1</sub> et un chat<sub>2</sub>, celui\*(-ci<sub>1</sub>) chasse  
 Each time that I see a dog and a cat, DEM-him\*(-HERE) chases  
 celui\*(-là<sub>2</sub>).  
 DEM-him\*(-THERE).

Finally, examples (13-14), which come from the French literature, show mixed uses of *-ci/-là* to express both abstract reference and contrast. In (13), the context helps resolve the referents targeted by *-ci* and *-là* (authors Racine and Corneille, respectively), as there does not seem to be a lexically encoded one-to-one mapping between {*-ci*, *-là*} and items such as {*the former*, *the latter*} in English. In (14) the referents picked up by *-ci* and *-là* remain totally underspecified.

- (13) Corneille<sub>1</sub> nous assujettit [...] à ses idées, Racine<sub>2</sub> se conforme aux nôtres;  
 Corneille us subjects [...] to his ideas, Racine REFL complies to ours;  
**celui-là**<sub>1</sub> peint les hommes comme ils devraient être,  
 DEM-him-THERE depicts the men as they should be,  
**celui-ci**<sub>2</sub> les peint tels qu’ils sont.  
 DEM-him-HERE them depicts as than-they are.  
 ‘Corneille subjects us to his ideas, Racine follows ours; the former depicts men as they should be, the latter, as they are.’ (Jean de La Bruyère, *Les Caractères*, 1696)

- (14) Ça tient, les options politiques, [...] parfois du hasard. [...] On a  
 It depends, the options political, [...] sometimes to chance. [...] One has  
 rencontré **celui-là** plutôt que **celui-ci**, et tout est joué.  
 met DEM-him-THERE rather than DEM-him-HERE, and all is played.  
 ‘Political options sometimes hinge on chance. We met this person rather than that person, and all is played out.’ (Alphonse Boudard, *Les combattants du petit bonheur*, 1977)

Turning to the restrictions slot, we assume it can host an NP as in (1), a strong pronoun as in (2) or a set of features with no overt exponent, as in (5-6). This last point is consistent with Ahn’s view and the intuition expressed by Kayne and Pollock (2010) that bare

*ce* constructions take a covert argument they dubbed “THING”. French crucially realizes these three options *transparently*, using the same *ce*-{*ci*, *là*} “wrapper” structure, unlike English, which uses some degree of suppletion in its DEM forms (e.g. *this/that*) or DEM pronouns (e.g. *these/those*). In the following section we will outline some limitations of the application of Ahn’s model to the French case, and sketch potential solutions.

### 3. Some puzzles for DEM structures at the syntax-semantics interface

#### 3.1 The need (or lack thereof) of an overt relation R

As outlined in Section 1, *ce*-NP is a standalone structure, while *ce*+*pro*<sub>strong</sub> structures require an overt realization of R, either in the form of *-ci/-là*, or in the form of a RC. This could be explained by the fact that the denotation of NPs is usually more specific than that of pronouns. For instance, the (still very generic) noun *man* denotes a set of human, masculine, atomic, adult individuals, while the pronoun closest to it, *he/him*, will in Ahn’s framework denote a set of human, masculine, atomic, but not necessarily adult, individuals. This difference in specificity between noun and pronouns might make the use of an overt R less useful to delineate the referent in the DEM+NP case as opposed to the DEM+*pro*<sub>strong</sub> cases, from a pragmatic perspective. More generally this suggests a division of labor between the restriction slot and the R slot, in terms of their semantic contribution.

#### 3.2 Differences in perceived animacy between *pro*<sub>strong</sub> and DEM+*pro*<sub>strong</sub>

As briefly mentioned in in Section 1 (example (2) in particular), DEM+*pro*<sub>strong</sub> (e.g. *celui*), contrary to its bare *pro*<sub>strong</sub> (e.g. *lui*), preferably denotes inanimate individuals. This is merely a preference, which, for instance, vanishes in cases of coordination, as noted by Cardinaletti and Starke (1999), suggesting that the contrast is not directly encoded in the syntax of those forms. Additionally, DEM+*pro*<sub>strong</sub> distributes just like a strong element w.r.t. topicalization (15a), short answers (15b), coordination (15c), object placement (15d), modification by adverbs (15e), reference to expletives (15f).

(15) Adapted from Cardinaletti and Starke (1999):

- a. {**Lui**, **Celui-ci**}, il est beau.  
{3.M.SG.STR, DEM-3.M.SG.STR-HERE}, 3.M.SG.WK is pretty.
- b. –Lequel est le plus beau? –{**Lui**, **Celui-ci**}.  
–Which is the most pretty? –{3.M.SG.STR, DEM-3.M.SG.STR-HERE}.
- c. **Celui-ci** et **celui-là** sont beaux.  
DEM-3.M.SG.STR-HERE and DEM-3.M.SG.STR-THERE are pretty.
- d. Marie trouve {**LUI**, **celui-ci**} beau.  
Marie finds {3.M.SG.STR, DEM-3.M.SG.STR-HERE} pretty.
- e. Marie aime {**lui**, **celui-ci**, \*il} aussi.  
Marie likes {3.M.SG.STR, DEM-3.M.SG.STR-HERE, \*3.M.SG.WK} too.
- f. {\***Lui**, \***celui-ci**} (il) pleut.  
{3.M.SG.STR, DEM-3.M.SG.STR-HERE} (3.M.SG.WK) rains.

We take this pattern as evidence that neither DEM+pro<sub>strong</sub> nor pro<sub>strong</sub> are lexically specified for [ $\pm$ animate], and that pro<sub>strong</sub> *acquires* the [+animate] specification by competition, resulting from two factors: (i) pro<sub>strong</sub> being structurally simpler than DEM+pro<sub>strong</sub>, and (ii) strong forms being empirically more likely to refer to humans. As a secondary consequence, DEM+pro<sub>strong</sub> ends up *preferably* denoting [-animate] entities.

### 3.3 The intricate distribution of *ce(cil/la)* w.r.t RCs and CPs

Examples (6) and (5) respectively showed that whenever *-cil/-là* marking is present in bare DEM structures, RC are banned while CPs are possible; and that whenever *-cil/-là* marking is *absent* the presence of either a RC or a CP becomes mandatory, the choice between the two being a function of the semantics of the main verb. This interaction is intriguing because it seems to contradict the idea that *-cil/-là* and RCs occupy different slots in the demonstrative structure and hence do not compete. Additionally, the role of CP within the demonstrative structure remains unaccounted for. We do not have a full solution to the pattern at stake, but put forth a few ideas in the next paragraph.

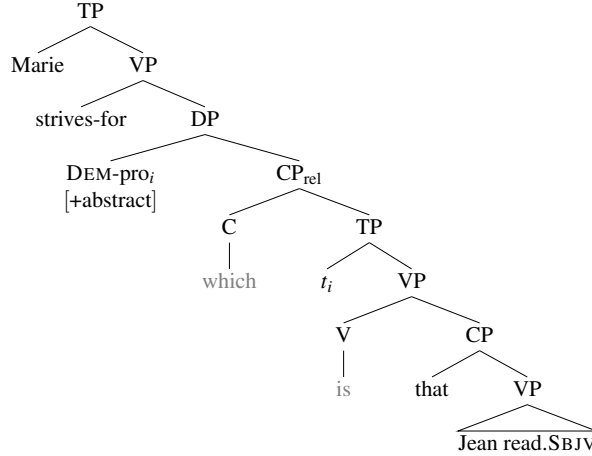
First, we want to suggest that the null pronoun combining with DEM (*ce*) in (5) and (6), may denote either a concrete [-animate] individual, or an abstract “individual with propositional content” in the sense of Moulton (2015). Which type of referent gets preferred depends partly on verbal selection, partly on the overall meaning of the sentence in context.<sup>7</sup> The former case (concrete inanimate referent) would apply to (5a)/(6a), and the latter case (abstract propositional referent), to (5b)/(6b). Given those assumptions, one might want to explain the pattern in (6a)/(6b), by positing that the CP in (6b) is in fact encapsulated within a covert predication RC (cf. (16)). This makes the CP “compatible” with the R slot and renders (6b) analog to (6a). The necessity of an RC in both structures was the topic of Section 3.1. Turning to the contrast (5a)/(5b), we suggest that (5b) results from extraposition, so that the demonstrative and the CP are coindexed. This is made possible by the presence of *-cil/-là*, used as anaphoric indices. This line of reasoning is illustrated in (17). Why (5a) disallows an extra non-restrictive RC remains however unclear, but might find a partial explanation in the featural underspecification of the demonstrative.

---

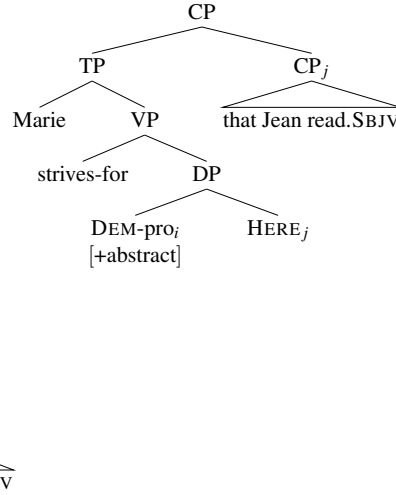
<sup>7</sup>Note that the second interpretation is overall preferred in the absence of strong contextual or syntactic cues, which again can be explained by appealing to some form of semantic competition between pronominal forms, and namely, the fact that (i) there is no overt pronoun referring to abstract entities like propositions or events in French, and (ii) concrete inanimate entities can already be referred to using DEM+pro<sub>strong</sub>.



- (16) Structure posited for (6b) (grayed heads remain unpronounced).



- (17) Structure posited for (5b) after extraposition of the CP.



#### 4. Conclusion

We showed how the French demonstrative paradigm is a fairly transparent illustration of Ahn (2022)’s unified account of demonstratives. More specifically, we identified the demonstrative morpheme *ce* with Ahn’s binary maximality operator (bi-sup), which can get restricted in French by a variety of elements, including NPs, strong pronouns, and null pronouns. We also argued that the markers *-ci/-là* instantiate the relation argument (R) passed to bi-sup; which explains why those markers compete with restrictive relative clauses. We then provided a “fused” semantics for those two markers, merging deictic and anaphoric functions. As a result, *-ci* and *-là* can be seen as higher-level “locators”, either in the physical space, or in the abstract domain of variables. It is worth noting that the French data presented here share some similarities with Afrikaans, as well as colloquial Swedish and Norwegian which also use *HERE* and *THERE* particles (Bernstein 1997, Leu 2007), although within potentially smaller paradigms.

Some aspects of French demonstratives however remain to be clarified and/or investigated. First, from the point of view of pure implementation, we did not fully explain how the *relative* character of proximal/distal contrasts could be accounted for based on the entries posited for *-ci* and *-là*,<sup>8</sup> nor did we fully spell out the kind of binding restrictions those markers introduce in the case of anaphoric indexation. Second, from a more empirical perspective, one might wonder why the distribution of *ce(ci)+CP* seems restricted to prepositional verbs such as *aspirer/viser à* (‘strive for’/‘aim at’), *consister/résider en* (‘consist/lie in’), *dériver/découler de* (‘derive/stems from’) etc. Is it simply that verbs taking propositional arguments have a tendency to combine with prepositions in the first place? Why is

<sup>8</sup>Recall that, whenever more than one demonstrative gets used in the context of deictic pointing, *-là* can be used across the board without *-ci*, but whenever a *-ci* gets introduced, one *-là* should preferably follow. This suggests that *-là* represents a deictic default, while *-ci* is expected to convey some form of deictic contrast.

that so? Another area of further investigation pertains to combinations between locative prepositions, such as *par* ('through') or *de* ('from'), and the *-ci/-là* markers. Are those constructions part of the demonstrative paradigm as well, or do they simply reuse the *-ci/-là* markers for slightly different purposes? Finally, we did not even touch upon the question of the free-relative reading of *ce que*, for which Ahn suggests DEM combines with no restriction, nor did we discuss the availability of subject bare *ce* in predicative sentences, discussed in (Kayne and Pollock 2010).

## References

- Ahn, Dorothy. 2022. Indirectly direct: An account of demonstratives and pointing. *Linguistics and Philosophy* 45:1345–1393. URL <https://doi.org/10.1007/s10988-022-09350-5>.
- Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102:87–113. URL <https://www.sciencedirect.com/science/article/pii/S0024384196000460>.
- Cardinaletti, Anna, and Michal Starke. 1999. *The typology of structural deficiency: A case study of the three classes of pronouns*, 145–234. Berlin, New York: De Gruyter Mouton. URL <https://doi.org/10.1515/9783110804010.145>.
- Kayne, Richard S., and Jean-Yves Pollock. 2010. Notes on French and English demonstratives. In *Linguistik aktuell/linguistics today*, 215–228. John Benjamins Publishing Company. URL <https://doi.org/10.1075/la.164.24kay>.
- Leu, Thomas. 2007. These HERE demonstratives. In *Proceedings of the 30th Annual Penn Linguistics Colloquium*, ed. by T Scheffler, J Tauberer, A Eilam, and L Mayol, volume 13, 141–154. University of Pennsylvania Working Papers in Linguistics.
- Moulton, Keir. 2015. CPs: Copies and compositionality. *Linguistic Inquiry* 46:305–342. URL <http://www.jstor.org/stable/43695680>.