Demonstrative Shift and Proximal Markedness

Tess Monks and Kathryn Davidson
Meaning & Modality Lab
Harvard Linguistics

Roadmap

Our focus

Despite a lot of work on demonstratives and definites within semantics and within the study of language change, there has been little focus on understanding how **distal** (English: *that*) and **proximal** (English: *this*) compare/contrast

We present an **experimental study** testing the acceptability of distal and proximal demonstratives across different contrastive contexts in both English and Italian, and propose a way to understand these results in light of known historical patterns. This work is in the spirit of motivating diachronic change patterns with synchronic experimental data (Zhang, Piñango & Deo, 2018).

Background on Grammaticalization Clines

Demonstratives are **"truly universal"** (Diessel, 2012: 12), since they appear in all languages (Himmelmann, 1997; Dixon, 2003)

As speakers use demonstratives in contexts where they are not strictly required for identification of a referent, their properties may diminish in importance and lose their integrality within the meaning of the demonstrative toward a variety of functional items (Epstein, 2002).

The path of grammaticalization from demonstrative to functional items has been widely studied (Wackernagel, 1924; Lyons, 1999; Selig, 1992; Greenberg, 1978; Lehmann, 2015 [1982]; Diessel, 1999)

Wackernagel (1928)

Second position reserved for certain particles or clitics in Indo-European

Adnominal demonstratives in second position lost emphasis and complexity because of this position's syntactic priming

Treatment of demonstratives as nominal markers began through their syntactic positioning

Lyons (1999: 161, 332)

Shift from demonstrative to definite article begins only in the contexts of a directly perceivable visual object or easily accessed referent. If such is the case, then deixis is unnecessary and thus lost as an essential component of meaning. Then, once demonstrative use is no longer combined with deixis, it loses this as a necessary part of the definition.

Greenberg's (1978) 'Cycle of the Definite Article'

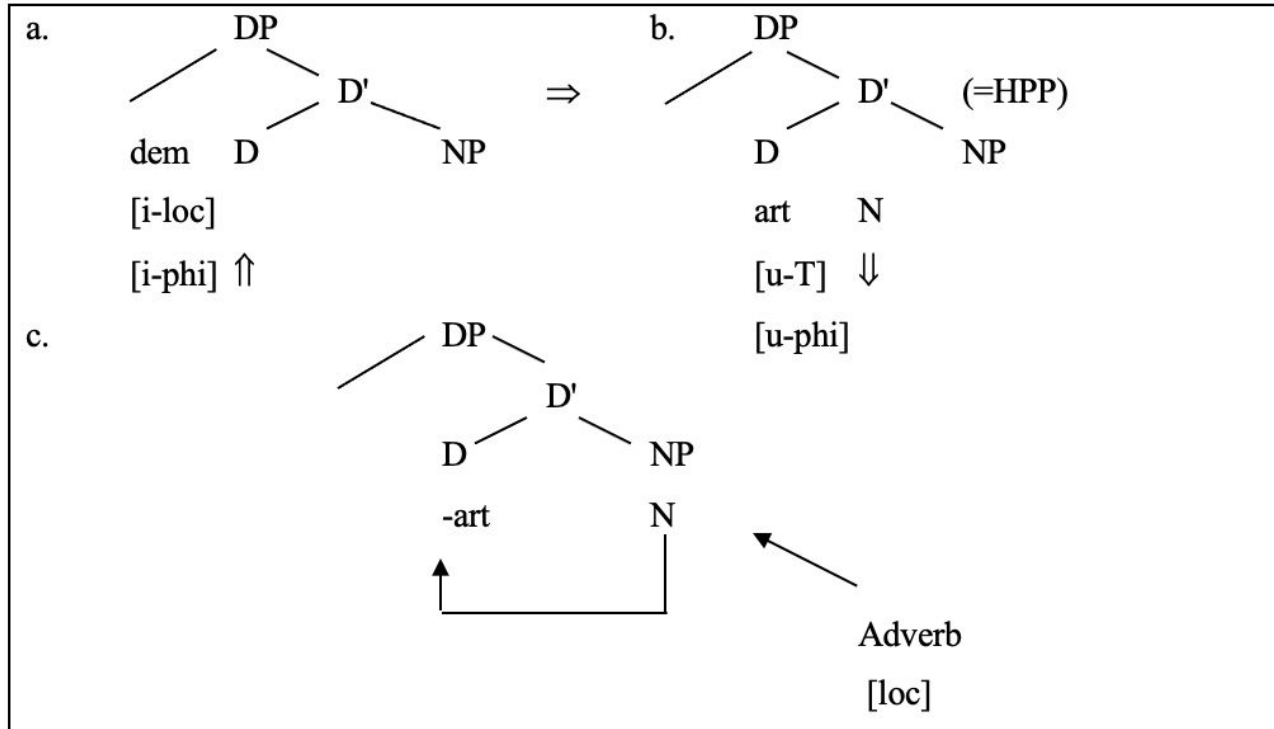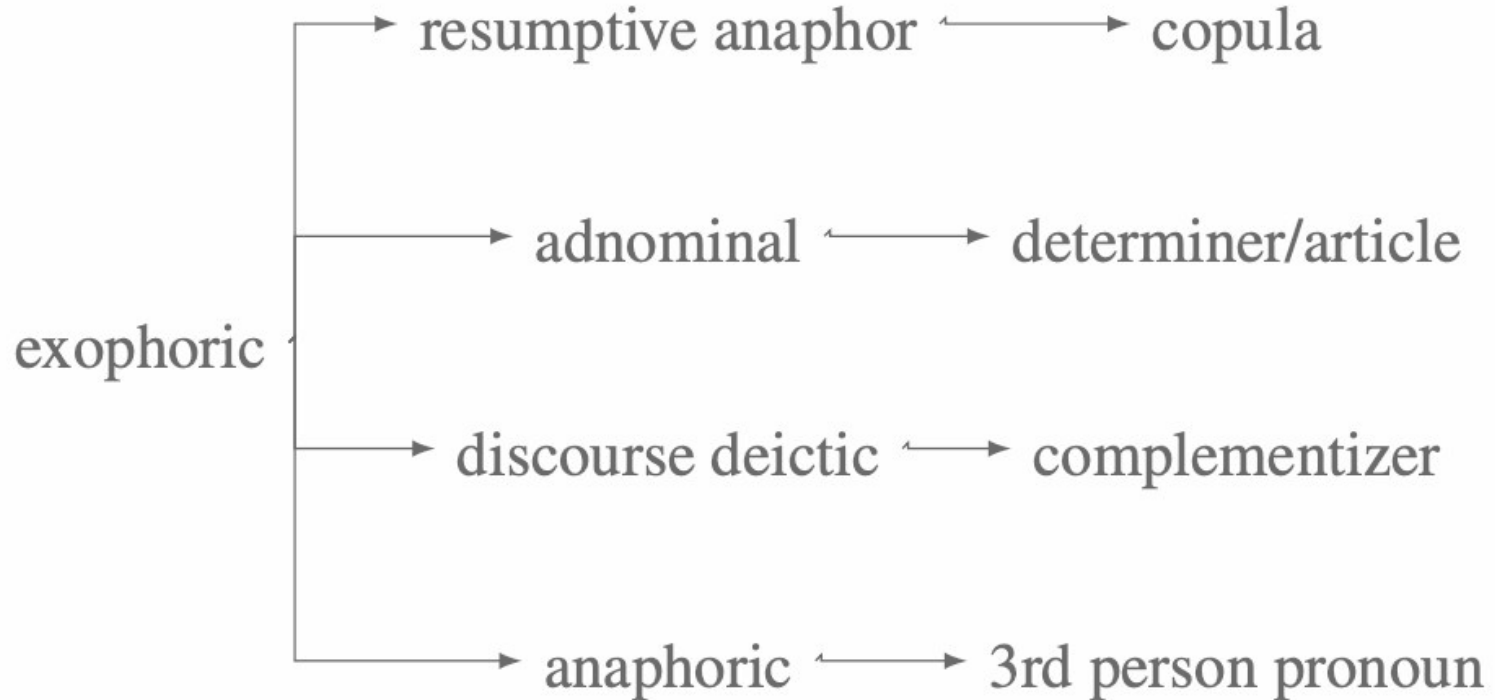| Stage 0		Stage I		Stage II		Stage III
Demonstrative	>	Definite Article	>	Specific Article	>	Noun Marker

van Gelderen's (2010) 'DP Cycle'

Figure 6.1: The DP Cycle ('dem' stands for demonstrative and 'art' for article)

Adapted from Diessel 1999

As for what happens before the first step of this cline:

"[T]here is **no evidence** from any language that demonstratives developed from a lexical source or any other source […]" (Diessel 1999: 150)

"Demonstratives are generally so old that their roots are **not etymologically analyzable.**" (Diessel 2006: 481)

Theories of Demonstrative Definition

There are many conflicting views on the definition of the demonstrative in step 1 of each of these processes of grammaticalization:

Descriptive explanation of demonstratives:

"[In defining the demonstrative,] the **delimitation from other possible references** is crucial. It is decisive that **other possible referents** exist in the common knowledge and that the designated referent can be clearly differentiated from these other ones by the hearer considering the common knowledge with the speaker, and the hearer can regard this difference as relevant." (Bisle-Müller, 1991: 80)

Diessel (2006: 469)

The two basic functions of demonstratives:

"1. First, they indicate the location of a referent relative to the deictic centre. 2. Second, they serve to coordinate the interlocutors' joint attentional focus."

Epstein (1993: 129)

A demonstrative is a linguistic point since it both identifies a referent in the immediate vicinity of the interlocutors and it focuses attention.

Logical Representations:

Elbourne (2009)

Definitions require situations with one unique satisfier, so a constrained domain is crucial for demonstratives. This definition is synchronic without a diachronic claim.

(81) $[\![\text{that}]\!]^{w,h,a,t} = \lambda x \,.\, \lambda f_{\langle e,sest\rangle} \,.\, \lambda g_{\langle se,st\rangle} \,.\, \lambda s \,.\, \iota z(f(x)(\lambda s'\,.\, z)(s) = 1 \ \&$
$g(\lambda s'\,.\, z)(s) = 1 \ \& \ \text{distal}(x, w, a, t))$

(82) $[\![\text{this}]\!]^{w,h,a,t} = \lambda x \,.\, \lambda f_{\langle e,sest\rangle} \,.\, \lambda g_{\langle se,st\rangle} \,.\, \lambda s \,.\, \iota z(f(x)(\lambda s'\,.\, z)(s) = 1 \ \&$
$g(\lambda s'\,.\, z)(s) = 1 \ \& \ \text{proximal}(x, w, a, t))$

(13) $[\![\text{the}]\!] = \lambda f_{\langle se,st\rangle} \,.\, \lambda s \,.\, \iota x f(\lambda s'\,.\, x)(s) = 1$

Patel-Grosz & Grosz (2017)

$$[\![\text{PER}]\!]^g = [\![(7a)]\!]^g = \iota x [[\![\mathbf{NP_n}]\!]^g(x)(s_r)]$$

$$[\![\text{DEM}]\!]^g = [\![(7b)]\!]^g = \iota x [[\![\mathbf{NP_n}]\!]^g(x)(s_r) \ \& \ x = g(1)]$$

Difference between personal pronoun and demonstrative is the requirement for an additional index argument that picks out the the demonstratum at that index. This definition does not differentiate between the distal and the proximal demonstrative and treats all locations as equal. Again, has no diachronic claim.

Ahn (2018)

(54) $[\![bi\text{-}sup]\!] = \lambda P.\ \lambda R.\ \iota x : \forall y\ [P(y) \wedge R(y) \leftrightarrow y \sqsubseteq x]$

R can denote deixis, relative clause, or anaphoric index. The degradation of R is what spurs extension of the demonstrative into myriad contexts, but Ahn's work does not interact with language change.
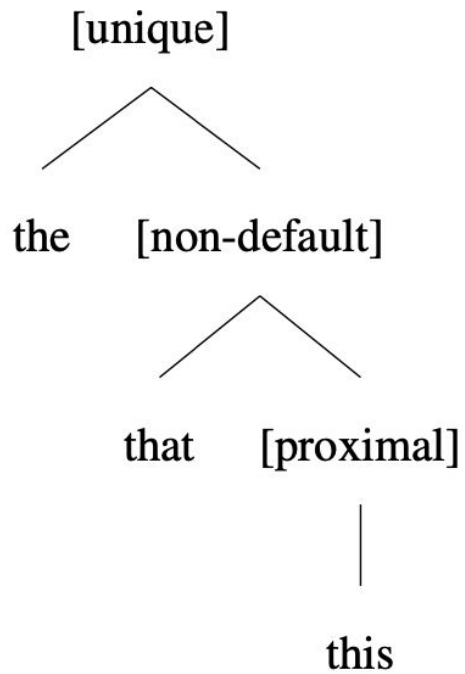
Wolter (2006)

$[[\mathit{that}_n]] = \lambda P.P(s_n)$ is a singleton set and s_n is non-default. If defined, denotes $\iota P(x)(s_n)$

$[[\mathit{this}_n]] = \lambda P.P(s_n)$ is a singleton set and s_n is non-default and $\iota P(x)(s_n)$ is proximal to the

speaker. If defined, denotes denotes $\iota P(x)(s_n)$

$[[\mathrm{the}_n]]: \lambda P.P(s_n)$ is a singleton set.
If defined, denotes $\iota x.P(x)(s_n)$

Wolter (2006), continued

[unique]
```
        [unique]
         /    \
       the   [non-default]
               /    \
            that  [proximal]
                      |
                     this
```

Even as some definitions treat the two terms as equally semantically complex or do not address the proximal/distal distinction at all, there is a tendency for most languages with grammaticalization clines from demonstrative to functional items to **grammaticalize the distal item** (Greenberg, 1978: 61; Himmelmann, 1997: 96-98).

Although this is true, and mostly focused in studies on Indo-European languages, it is a cross-linguistic phenomenon regardless of language family (Enfield, 2003; Piwek, 2008; Ferrazzano, 2013). It is not a rule, but rather a tendency.

Options for treatment of the distinction of proximal/distal in the formal semantic representation of demonstratives:

- + proximal or + distal as an explicit part of the definition

- [+/- proximal] or [+/- distal] as a feature system

- Not included in the essence of a demonstrative at all

- Purely presuppositional content

- [+ proximal] and unmarked for location

Experimental Question and Process

The tendency for languages to grammaticalize the distal demonstrative in binary demonstrative systems suggests an **asymmetry** between the demonstratives such that the distal is less marked than the proximal. Since there are more contexts in which a semantically more simplex form may extend, the distal demonstrative is semantically simplex in relation to the proximal by this reasoning.

In English, the distal demonstrative "that" grammaticalizes into an anaphoric pronoun, complementizer, and adnominal determiner (Diessel, 1999).

We hypothesize that the **less marked demonstrative** is more available to refer to more elements, regardless of their location in space with respect to the speaker. This extension makes available the less marked form for grammaticalization.
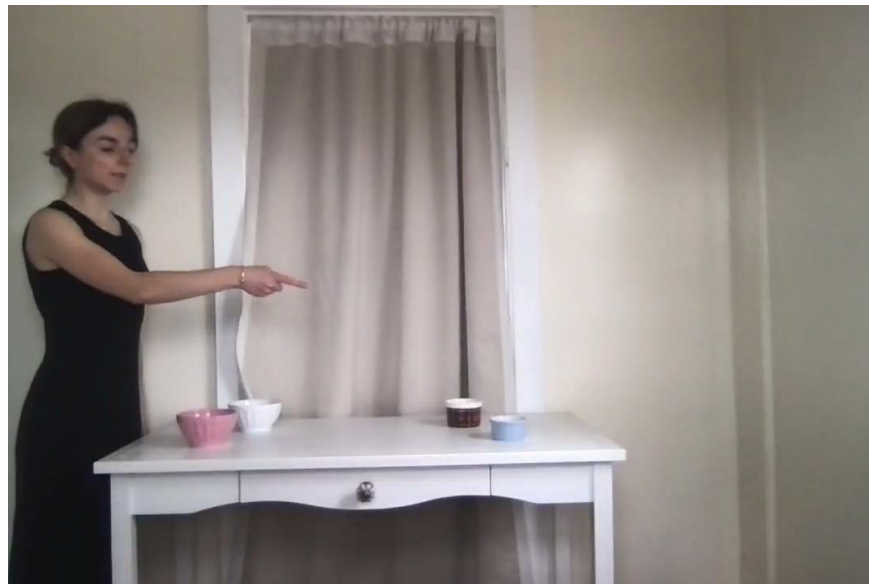
Experimental Format

We recruited 70 participants for the English study through Prolific Academic and filtered for self-identified native English speakers. We ran a total of two surveys, each with 35 participants and 18 questions per survey with every commutation of pronoun and location assessed. The questions within each survey were randomized and split across two surveys for counterbalancing.

Each question consisted of a short video prompt and a slider bar acceptability judgment.
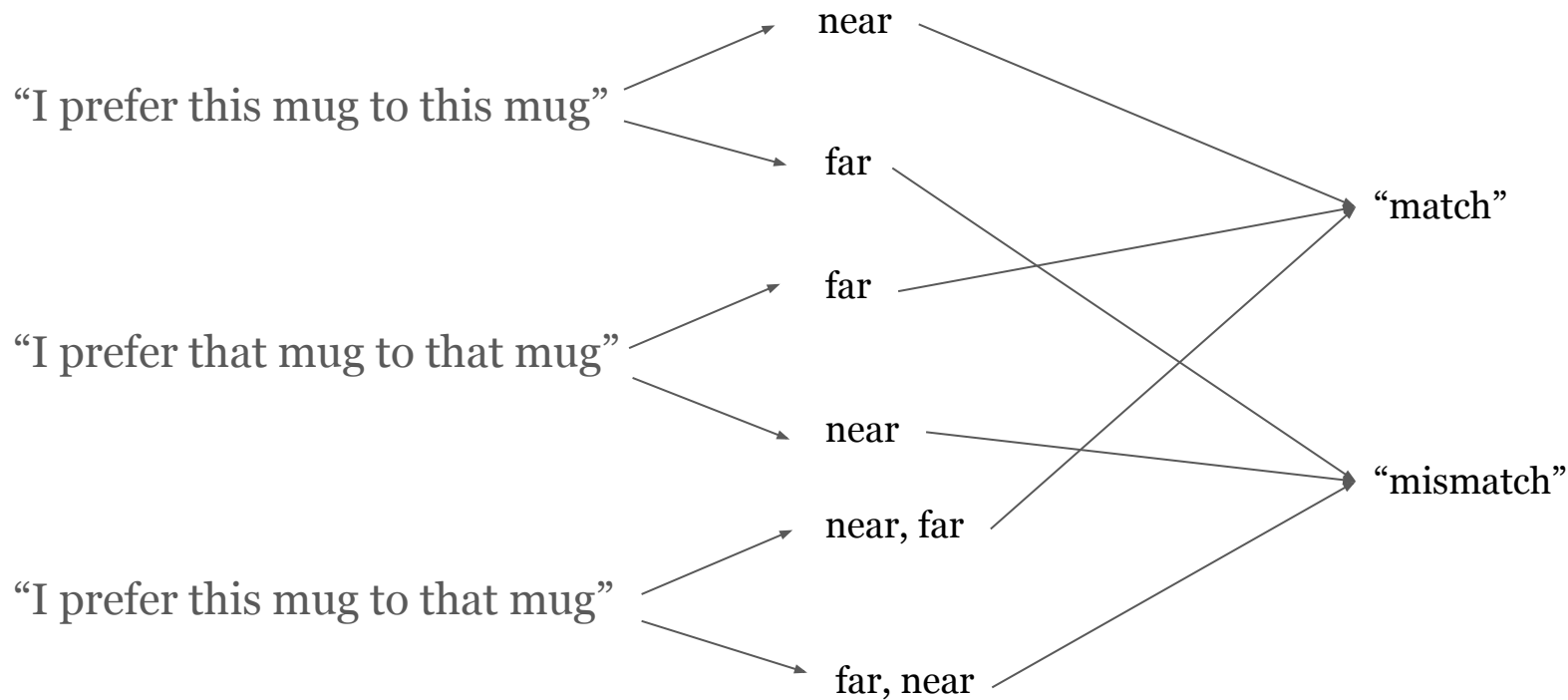
Please play the video below, then rate the acceptability of the sentence in the video.

Least Acceptable   Most Acceptable

Rating:

"I prefer this mug to this mug" → near, far

near → "match"

far → "mismatch"

"I prefer that mug to that mug" → far, near

far → "match"

near → "mismatch"

"I prefer this mug to that mug" → near, far ; far, near

near, far → "mismatch"

far, near → "match"

NPs included: books, pens, bowls, mugs, glasses, shoes (6 NPs * 6 commutations = 36 sentences)

	Location Match	Location Mismatch
This	71.703	53.158
That	73.394	61.879

Acceptability
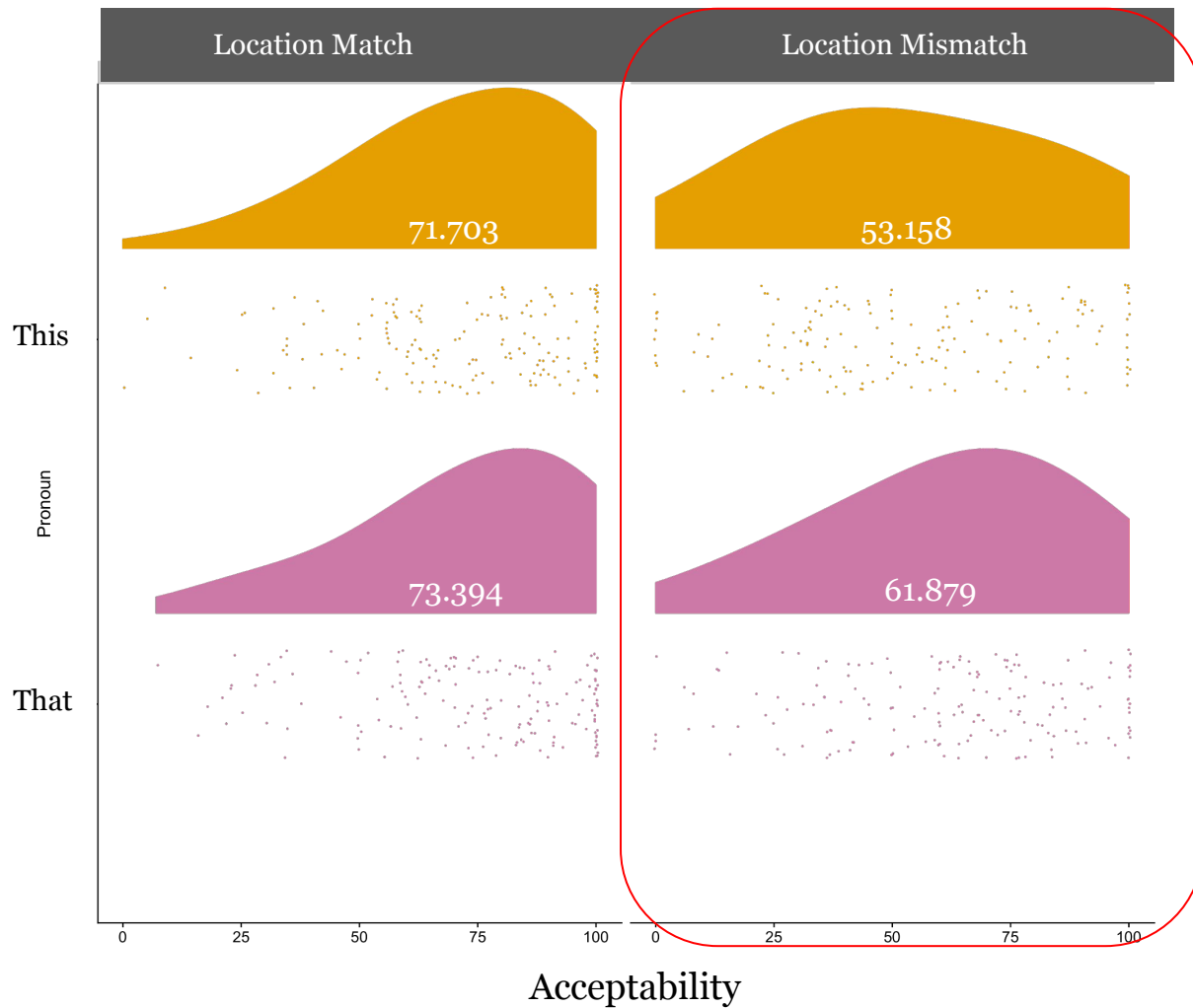
Mixed effects logistic regression in R

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	73.394	1.9773	7.128	< 2e-16 ***
PronounThis	-20.236	2.796	-7.239	1.27e-12 ***
LocationNear	-11.515	2.796	-4.119	4.29e-05 ***
PronounThis:LocationNear	30.061	3.954	7.603	1.00e-13 ***

The English data supports an **asymmetry in the markedness** of proximal and distal demonstratives. The source of grammaticalization is this markedness in the demonstrative systems such that the distal demonstrative is more available to shift from lexical to functional item.

In using this data to define the demonstrative, it is necessary to convey the asymmetry in the logical forms of the demonstratives such that the **less semantically complex form is for the distal.**

As another language with a bipartite system from another language family, namely Romance rather than Germanic, Italian proves an important extension of this line of inquiry.

Italian's synchronic demonstratives do not function as complementizers, definite articles, copulas, or other fully grammaticalized items. Instead, they function mainly as contrastive exophors. The synchronic definite articles (*il, la,* etc.) and the third person pronouns (*lo, li, la, le*) are grammaticalized forms of Latin distal demonstrative *ille*.

Latin's distal demonstrative *ille* competed with *ipse* (SELF) to grammaticalize along the cline toward the definite article (Carlier & De Mulder, 2010). To be able to compete, *ille* had to be semantically simple as compared to *ipse*. Since *ille* was semantically simple enough to refer to more objects, it competed with and eventually beat out *ipse* in extending along the cline. *Ille* solely grammaticalized, either by phonetic (Banniard, 2001), socio-linguistic (Aebischer, 1948), or structural primes (Selig, 1992).

Since *ipse* and *ille* coexisted, they were semantically simple to the same degree as compared to other demonstratives or morphs accessible to grammaticalize in this direction.

Italian's diachrony proves that the grammaticalization cline from distal demonstrative to grammatical items is readily available in the language.

We predict that this cycle repeats when a new lexical item is adopted into the place of the exophoric demonstrative. Therefore, *questa* (proximal) and *quella* (distal) should at some point resemble the data from English.
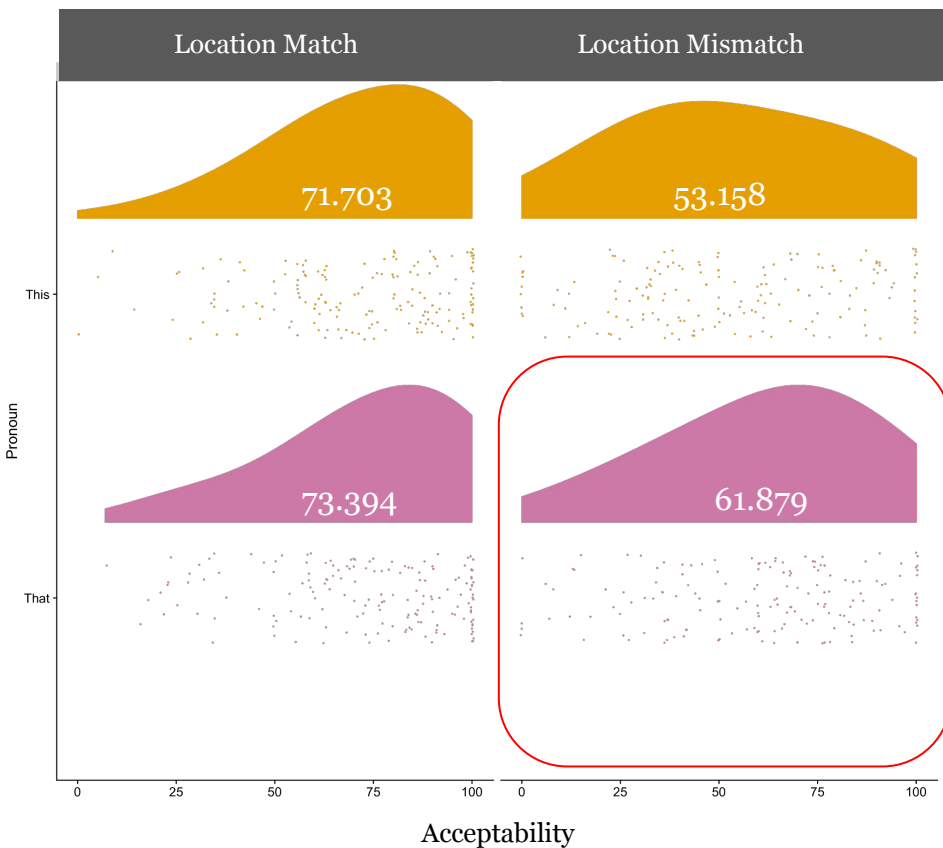
Acceptability
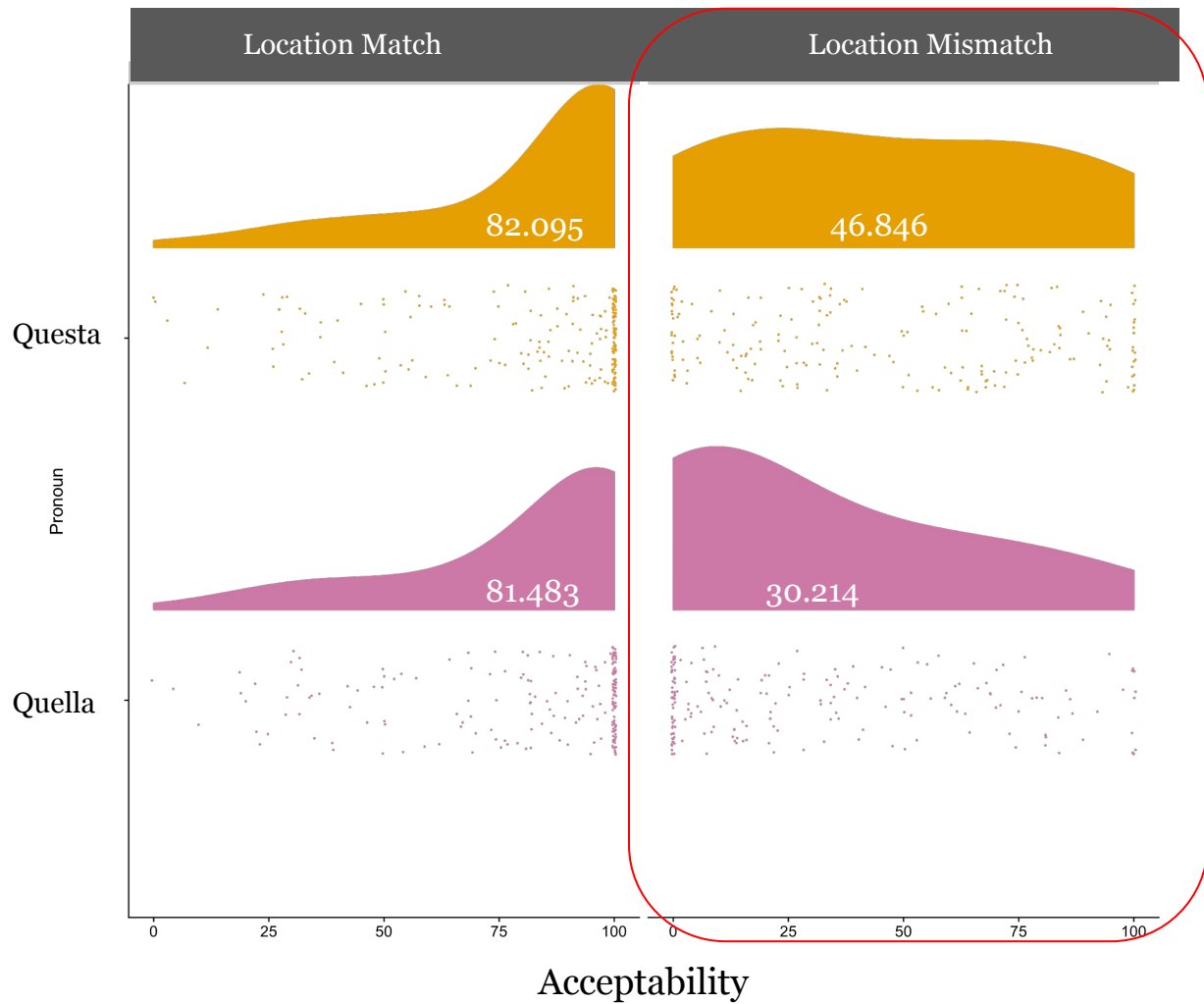
Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	81.483	2.091	38.97	<2e-16 ***
PronounQuesta	-34.637	2.957	-11.71	<2e-16 ***
LocationProximal	-51.269	2.957	-17.34	<2e-16 ***
PronounQuesta:LocationProximal	86.517	4.182	20.69	<2e-16 ***

English

Location Match	Location Mismatch
71.703	53.158
73.394	61.879

This

That

Pronoun

Acceptability

Italian

Location Match	Location Mismatch
82.095	46.846
81.483	30.214

Questa

Quella

Pronoun

Acceptability
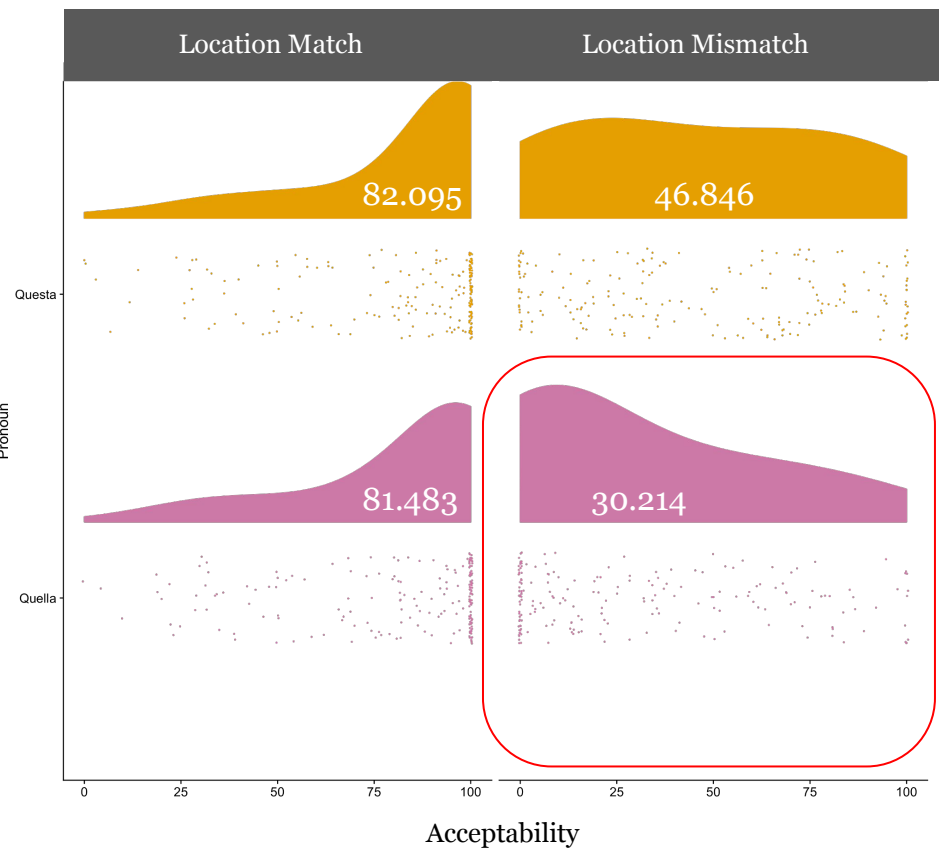
The Italian data, though in opposition to the English data, may serve as an important step in **diagnosing the position along the grammaticalization cline** a certain morph may lie at a synchronic moment. Since Italian has different lexemes for the definite article (*il, la*), the complementizer (*che*), the pronoun (*il, la*), and does not have a non-verbal copula formed from the demonstrative root, it may be the case that the purely deictic form of the demonstrative is available.

The Italian demonstratives may not yet be extended to more contexts to allow for greater use in more contexts.

The proximal may be slightly less marked in this survey than the distal because of language-specific considerations for the proximal space. Speakers in English and Italian may have different diameters of proximity such that the objects in the videos actually were only proximal to the Italian participants. English participants used the distal more readily in all contexts, even where the proximal was allowed. Italian participants dispreferred the distal in the proximal context more than the English participants.

The Italian data may prove an important clue into diagnostic approaches for grammaticalization clines. Synchronic snapshots of languages may show the stage of a language in the cycle towards a grammatical object.
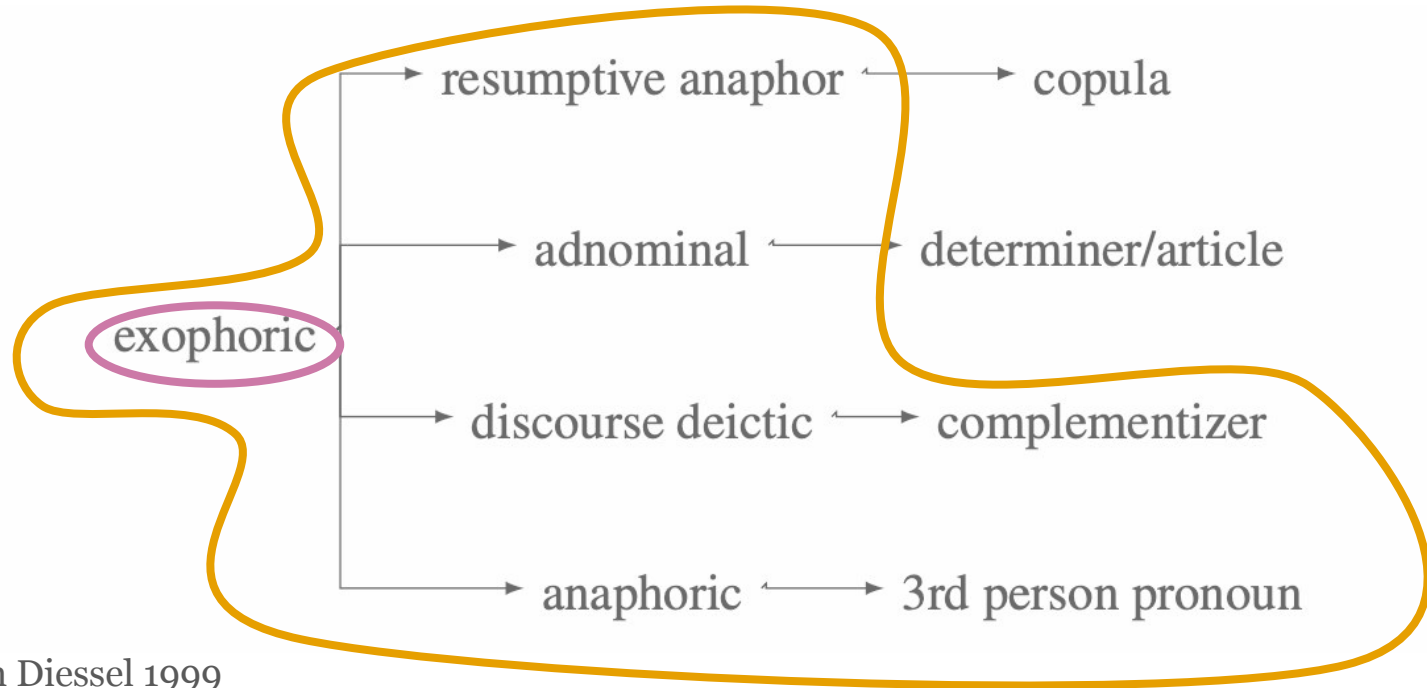
The Diachronic Role of Semantic Complexity

Logical Forms

"Marked" for proximal

$$[[this]] = \lambda P . \iota x P_{w0}(x) \wedge at_{w0}(x)(l_n) \wedge speaker\ pointing\ at_{w0}(l_n) \wedge \overbrace{close_{w0}(l_n)}$$

$$[[that]] = \lambda P . \iota x P_{w0}(x) \wedge at_{w0}(x)(l_n) \wedge speaker\ pointing\ at_{w0}(l_n)$$

The empirical asymmetry supports thinking of the change in the demonstrative from its strong deictic meaning to the definite article meaning as involving a gradual loss of features.

We adopt a presuppositional analysis of the definite article as requiring uniqueness and, if defined, picking out the one unique referent:

$$\lambda P.\iota x P_{w0}(x)$$

As movement from "that" to the definite article, we first lose the pointing element of the demonstrative.

$$\lambda P.\iota x P_{w0}(x) \wedge at_{w0}(x)(l_n) \wedge speaker \ pointing \ at_{w0}(l_n)$$

The second element is either locational in a situational or discourse sense ("that theory"), such that the anaphoric use relies on the closeness or distance of the referent in the discourse to determine which of the demonstratives to use. Since the demonstrative "that" does not require a specific location, it can also be used in the discourse anaphoric sense for both close and far antecedents.

$$\lambda P.\iota x P_{w0}(x) \wedge at_{w0}(x)(l_n)$$

$$[[this]] = \lambda P. \iota x P_{w0}(x) \wedge at_{w0}(x)(l_n) \wedge speaker\ pointing\ at_{w0}(l_n) \wedge \boxed{close_{w0}(l_n)}$$

$$[[that]] = \lambda P. \iota x P_{w0}(x) \wedge at_{w0}(x)(l_n) \wedge speaker\ pointing\ at_{w0}(l_n)$$

$$\lambda P. \iota x P_{w0}(x) \wedge at_{w0}(x)(l_n)$$

$$\lambda P. \iota x P_{w0}(x)$$

Since the loss of these elements leaves the demonstrative "that" identical to our definition of the definite article, it is then apparent that the distal, as the less marked version of the demonstrative, is the available demonstrative to grammaticalize along the cline.

In languages with a tripartite system, as Turkish or Spanish, it is generally the **middle distance** that grammaticalizes toward functional items. Not only are these elements less specified as proximal or far distal, they also have more uses in the discourse as they can refer to many salient items (Ferrazzano, 2013).

Conclusions:

The distal demonstrative is the most frequent candidate for grammaticalization along the cline across languages (Heine and Kuteva 2002, Greenberg 1978, Lyons 1999, Lehmann 2015, Diessel 1999) because of an asymmetry in markedness between the distal and the proximal demonstratives.

Since the distal is less marked for distance, the process of adopting new meanings and extending its functional uses diminishes the strength of deixis and locational requirements for demonstrata towards its use as a definite article.

Data collected from surveys of English and Italian speakers when presented with an acceptability judgment task supports this conclusion and provides an empirical basis for defining the demonstratives in an asymmetric manner.

Next Steps:

Extending this experimental approach to a language with a tripartite system, such as Turkish, would help illuminate the markedness distinction, if any, within a language where the middle demonstrative is targeted for grammaticalization.

Defining the process of the demonstrative's grammaticalization toward the copula, pronoun, and complementizer would also help illustrate the latent elements within the demonstrative's lexical form that allow it to extend into such roles.

References:

Aebischer, Paul. 1948. *Estudios de toponimia y lexicografía románica*. Barcelona: Escuela de filologiá.

Ahn, Dorothy. 2018. THAT thesis: A competition-based mechanism for anaphoric expressions. Doctoral dissertation, Harvard University.

Banniard, Michel. 2001. "Causes et rythmes du changement langagier en Occident Latin (IIIe-VIIIe s.)." Travaux Neuchatelois de Linguistique (Tranel), vol. 34-35. 85-99.

Carlier, A. & W. De Mulder. 2010. "The emergence of the definite article in Late Latin : ille in competition with ipse." In : H. Cuykens, K. Davidse & L. Van de Lanotte Eds, *Subjectification, intersubjectification and grammaticalization*. The Hague: Mouton De Gruyter. 241-275.

Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins Publishing Company.

Dixon, R. M. W. 2003. "Demonstratives: A cross-linguistic typology." Studies in Language, 27(1). 61–112.

Enfield, 2003. "Demonstratives in Space and Interactions: Data from Lao Speakers and Implications for Semantic Analysis." Language 79(1). 82-117.

Epstein, Richard. 2002. "The definite article, accessibility, and the construction of discourse referents." CognitiveLinguistics 12 (4). 333–378

Ferrazzano, Lisa Reisig. 2013. "Demonstratives in Motion: The Grammaticalization of Demonstratives as a Window into Synchronic Phenomena." PhD dissertation. The City University of New York.

Greenberg, Joseph Harold. 1978. "How does a language acquire gender markers?" *Universals of Human Language. Volume 3: Word Structure*, ed. By Joseph Harold Greenberg, Charles A. Ferguson, and Edith Moravcsik, 47-81. Stanford, CA: Stanford University Press.

Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9783110929621

Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer. https://doi.org/10.1515/9783110929621.

Lehmann, Christian. 2015 [1982]. *Thoughts on grammaticalization*. Berlin: Language Science Press. https://doi.org/10.26530/OAPEN_603353.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: University Press. https://doi.org/10.1017/CBO9780511605789.

Patel-Grosz, Pritty & Patrick Grosz. 2017. "Revisiting Pronominal Typology." Linguistic Inquiry 18 (2). 259-267.

Piwek, Paul et al. 2008. "'Proximal' and 'distal' in language and cognition: Evidence from deictic demonstratives in Dutch." Journal of Pragmatics 40. 694-718.

Selig, Maria. 1992. *Die Entwicklung der Nominaldeterminanten im Spätlatein. Romanischer Sprachwandel und lateinische Schriftlichkeit*. Tübingen: Narr.

Wackernagel, Jacob. 1928. *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinsich und Deutsch,* 2. edition. Basel.

Wolter, Lynsey. That's That: The Semantics and Pragmatics of Demonstrative Noun Phrases. Doctoral dissertation, UCSC.

Zhang, Muye, Maria Mercedes Piñango & Ashwini Deo. 2018. Real-time roots of meaning change: Electrophysiology reveals the contextual-modulation processing basis of synchronic variation in the location-possession domain. *CogSci:2019.*

Questions?

Lehmann, 2015 [1982]

| Demonstrative determiner | → | Weakly demonstrative definite determiner | → | Definite article | → | Affixal article | → | Noun marker |

Harris (1978: 67-68)

A demonstrative's purpose is "to *specify* one particular item (as opposed to an indefinite number of alternatives) and to mark the proximity of that item, whether that proximity be spatial, temporal, or simply psychological…; a definite article, on the other hand, when fulfilling its original semantic function…, has only the first of these two functions, namely to *specify*… one particular referent as opposed to all others."