Presuppositionless proxies of politeness:
An (eventually) Optimality-Theoretic account

by

Ruoan Wang

M.A. in Linguistics
Queen Mary University of London, 2018

Submitted to the Department of Linguistics and Philosophy in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2025

Copyright 2025 Ruoan Wang. All Rights Reserved.
The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by ..... Ruoan Wang
Department of Linguistics and Philosophy
August 2025

Certified by ..... Adam Albright
Professor of Linguistics
Thesis chair

Certified by ..... Kai von Fintel
Andrew W. Mellon Professor of Linguistics
Thesis advisor

Certified by ..... Norvin Richards
Dean's Distinguished Professor of Linguistics
Thesis advisor

Accepted by ..... Danny Fox
Anshen-Chomsky Professor of Language & Thought
Linguistics Section Head

Presuppositionless proxies of politeness:
An (eventually) Optimality-Theoretic account

by

Ruoan Wang

Abstract

Recent research has seen growing interest in the nature of *social meaning*, "the constellation of qualities and properties that linguistic forms convey about the social identity of language users" (Beltrama 2020, see also Eckert 2008). This dissertation concerns those linguistic forms termed *polite pronouns*. They convey one particular aspect of social identity: the social distance between language users, where a large social distance mandates the expression of politeness.

Most existing work on polite pronouns has modeled polite meaning via dedicated grammatical mechanisms, such as a dedicated feature ([HON], Ackema & Neeleman 2016) or a dedicated dimension of meaning (McCready 2019). This dissertation shows that such dedicated mechanisms are superfluous, as *existing* grammatical mechanisms can be appropriated to describe and explain both the form and meaning of polite pronouns.

Building on Wang (2023), I use a purpose-built typological sample of polite pronouns in >220 genetically and geographically diverse languages to establish the shapes of polite pronouns, showing that polite pronouns are obtained by asymmetrical recruitment of existing ϕ-featural values. Furthermore, the shapes of polite pronouns converge precisely with the shapes of semantic defaults, those ϕ-featural values underspecified in meaning, which must emerge when the context provides little or no information about number or person. To explain this convergence, I use the notion of negative politeness: respecting an interlocutor's right to be unimpeded (Brown & Levinson 1987). I argue that semantic defaults and polite pronouns are morphologically identical because they are *pragmatically* identical: underspecification makes them well-suited to be avoidance mechanisms in the service of negative politeness. This captures the intuition that avoidance behaviors are a core component of expressing politeness. Specifically, polite pronouns enable speakers to avoid making presumptions about aspects of the interlocutor. Hence, presuppositionlessness, proxies, and politeness are what this dissertation is all about.

I implement this intuition in Optimality-Theoretic terms, where polite grammars are defined by an outranking of Markedness (which mandates speakers to avoid presupposition-rich forms) over Faithfulness (which mandates speakers to be as informative as possible). The resulting system is restrictive but powerful, delivering a factorial typology which exactly mirrors the recruitment asymmetries. With no stipulations specific to polite pronouns, the account is furthermore able to make concrete predictions about politeness phenomena beyond the ϕ-featural domain.

Happily, the resulting account is as lean as can be. Its main innovations are methodological and theoretical. Methodologically, the investigation is led by a large cross-linguistic sample, and the additional consideration of use conditions alongside morphological exponence. Theoretically, recruitment affords us with a fresh look at how the universal inventory of features is organized. This dissertation shows that recruitment is not only *viable* as a meta-grammatical operation, but even *desirable* for reasons of economy or computational efficiency.

Acknowledgements

I couldn't have written this monstrosity without the support of my advisors, who were all extremely generous in many ways. I hope I will have the opportunity to pay it forward in the future, because paying it backward to my advisors would not humanely possible. They were frighteningly encouraging of my ideas when those ideas were fresh out of the gob, patiently listening when those ideas were crazy, and of course, they read and re-read everything. A writer will probably never have this kind of editorial expertise at hand again, and I hope I have not taken advantage of them. Also a big thank you to Nina Haslinger, Viola Schmitt, Daniel Harbour, my Family, and the MIT Linguistics staff.

I couldn't have even gotten to the stage of monstrosity-writing without the support of the following mixx, who have come through for me so many, many, manymany times. I wouldn't be afloat without them and I wouldn't know why I am here (specifically, not existentially). When I suffer a small death from graduating, it is because we will have left each other's orbit, and it is my sincerest hope that we call on each other as phones do. They are: Keri's Gold's Chili Sauce (for indulging my sweet tooth and struggle tooth, and for allowing me to indulge theirs on rare occasions); Ally of Hèdgéhog and Mousie (for many shared giggles and band-aids, and for showing me how to be strong); Gourmet Judiciary (for their generosity and wisdom, offered casually and incisively, which really should be used to write novels); Honey Bunny (for the crucial education they provided in all things, especially in bunniness and foxiness, even from afar); The Soprano (for shooting straight to the heart of things); ze Graf (for having faith in all my abilities); ze Oma (for discovering happiness as a source of renewable energy—and sharing it); Leay (for their kindness, which is always issued immediately and magnanimously); B-ana-na (for their inspiring resilience); Bellybutton Hussy (for the fount of credit and goodwill extended and extended again and again when my own has dried up).

Out of vanity I would also like to thank Devin Townsend, Andrea Dworkin, Amia Srinivasan, Chino Moreno, Michael Cunningham, Sally Rooney, Franz Liszt, Rachmaninov, Jade Puget, Davey Havok, Dave Gahan, Martin Gore, Jenny Offill, Her Thighnesses, CN, MM, and HH, for reminding me to be brave.

To all those who rendered their ears

[Kinoko-san] also likes to say, "I'll just borrow your ear now, if I may." Sometimes I want to tell her, "An ear isn't the sort of thing you just borrow from other people whenever you feel like it," but then I swallow hard and hold my peace. I bite my nails and tell myself, "All right, then. If she wants an ear, I'll give her an ear. An ear isn't something you have less of just because you lend it out. Let's say you slice it off with one slash of a razor-fine! Then just wind a bandage around what's left." My speech, unlike Kinoko-san's, is becoming less elegant by the day.

When approaching another person, it's immature to begin by saying, "Um..." or "Uh..." On the other hand, "You know" is a bit too self-important, "Excuse me" is too formal, "By the way" too inefficient, and "Now then" too brisk. When you consider the options, Kinoko-san's ear-borrowing isn't so bad after all. Having heard it so often, I'm getting used to the phrase. You can get used to almost any phrase if you hear it every day.

But now Kinoko-san has come up with more and more radical variants. First she changed the phrase to "Rend me your ear," then to "render" and even "surrender." Since she's been growing more elegant daily, it isn't surprising to hear her append an extra syllable, or shift the consonants in one direction or another. Still, it makes my chest clench up to think of rendering my ears.

Yoko Tawada, *Spores*

Contents

1 Introduction **8**

2 The Journey Starts: Big Politeness **12**

2.1 Single-tier politeness systems . . . 16

2.1.1 Honorific uses of number . . . 16

2.1.2 Honorific uses of person . . . 20

2.1.2.1 Combinations of values . . . 22

2.1.3 Honorific uses of inclusive . . . 25

2.1.4 Distinct proxands of politeness . . . 26

2.1.4.1 Politeness directed to a third person . . . 27

2.1.4.2 Politeness/Humbling directed at the first person . . . 30

2.1.4.3 Politeness directed towards a group . . . 32

2.1.5 The negative evidence . . . 33

2.1.6 Recruits in free variation . . . 34

2.2 Multiple-tier politeness systems . . . 37

2.2.1 Tiered systems: An introduction . . . 38

2.2.2 Dravidian . . . 39

2.2.3 The specter of Indo-Aryan . . . 40

2.2.3.1 Eastern Indo-Aryan . . . 40

2.2.3.2 Central, Northern, Northwestern, and Western Indo-Aryan . . . 43

2.3 Moving on: Explicanda and Desiderata . . . 45

2.3.1 Explicanda: asymmetries and sub-asymmetries . . . 45

2.3.2 Desiderata . . . 45

3 On the way, a kindred spirit Semantic Defaults **47**

3.1 Social reflexes of politeness: Face and indirectness . . . 49

3.1.1 Face and face-work: Goffman . . . 49

3.1.2 Onto language: The facework of politeness . . . 50

3.1.3 “Indirectness” . . . 52

3.2 Diagnosing semantic defaults . . . 52

3.2.1 Manners of markednesses 54
3.2.2 Defaults in the category of person: When does a third person not quack like a third person? 56
3.2.2.1 Interim summary for person: Third person does not quack like a third person when it is person-neutral 60
3.2.3 Defaults in the category of number: When does a plural not quack like a plural? 60
3.2.3.1 Nominal number 61
3.2.3.2 Pronominal number 65
3.2.3.3 Interim summary for number: Plural does not quack like a plural when it is number-neutral 67
3.2.4 Shapes of arbitrary reference: Many quacks given 67
3.2.4.1 Hallmarks of impersonality 67
3.2.4.2 Routes to impersonality 68
3.2.5 Summary of default feature values 71
3.3 Diagnosing semantic defaults, cross-linguistically 73
3.3.1 Typological defaults for number and person 74
3.3.1.1 3SG as default 74
3.3.1.2 3PL as default 74
3.3.1.3 3SG, 3PL as defaults 75
3.3.1.4 Dedicated impersonal forms as defaults 78
3.3.1.5 Third persons as imposters 80
3.3.2 Typological defaults for clusivity 80
3.3.2.1 When does the inclusive not quack like an inclusive? When it quacks like an impersonal 81
3.3.2.2 Inclusives as imposters 87
3.3.2.3 A diachronic affinity between inclusive and impersonal pronouns: Indo-Aryan *APNE* 89

4 Privativity Plus **92**
4.1 Linguistic reflexes of indirectness: underspecification via privative features 94
4.1.1 The ontology and its LF Shadow 96
4.1.2 Privative ϕ-featural action on the lattice 99
4.1.2.1 Person: Restricting the ϕ-lattice with π features 99
4.1.2.2 Number: Restricting the ϕ-lattice with ω features 102
4.2 Conventional usages of pronouns: Further restricting the ϕ-lattice via strengthening 104
4.2.1 Impersonals: Wanton non-restriction of the ϕ-lattice 108
4.2.2 Inclusives 109
4.2.2.1 The inclusive: impersonal and conventional interpretations . . . 109
4.2.2.2 Existing analyses of the inclusive 111

4.2.2.3 A small intervention . . . 113
4.2.2.4 A tentative proposal for the inclusive: Two steps of strengthening 114
4.2.3 PF shadows . . . 116
4.3 Default usages of pronouns: Strengthening has defaulted . . . 118
4.4 Polite usages of pronouns: Strengthening has also defaulted . . . 122
4.4.1 The gist: Redressive efforts in the face of Taboo . . . 122
4.4.2 Weak forms as the output of Taboo . . . 125
4.4.2.1 Deriving polite address . . . 125
4.4.2.2 Deriving polite reference . . . 132
4.4.2.3 Deriving polite/humble self-reference . . . 134

5 The Final Analysis 136
5.1 Optimality Theory: Polite and non-polite grammars . . . 137
5.1.1 Deriving proxies of singular polite self-reference . . . 140
5.1.1.1 Predictions *vs.* reality for singular polite self-reference . . . 149
5.1.2 Deriving proxies of singular polite address . . . 150
5.1.2.1 Predictions *vs.* reality for singular polite address . . . 155
5.1.3 Deriving proxies of singular polite reference . . . 156
5.1.3.1 Predictions *vs.* reality for singular polite reference . . . 159
5.1.4 Deriving proxies of plural polite address . . . 160
5.1.4.1 Predictions *vs.* reality for plural polite address . . . 162
5.1.5 Deriving politeness clines . . . 163
5.1.6 Deriving sentential politeness clines: a sketch for mixed agreement . . . 164
5.1.7 Summary . . . 169
5.2 Other ϕ-featural categories: Gaps? . . . 170
5.2.1 Gender . . . 170
5.2.2 Spatial deixis . . . 172
5.3 Comparison to other analyses . . . 173
5.3.1 A dedicated grammatical feature for honorification . . . 173
5.3.2 Typological inadequacy . . . 175
5.3.3 An unusual feature, an unusual operation . . . 177

6 Conclusion 181

Chapter 1

Introduction

This dissertation poses a basic question: How can speakers say precisely what they mean when their language does not provide the relevant grammatical distinctions? The answer it poses is simple: recruit existing forms and make them do double duty. More specifically, this dissertation is about recruitment that expresses distinctions of a social nature: the social relationship that a speaker is situated relative to their interlocutor. It is crucially assumed that the core grammar itself does not encode such distinctions in the forms of a dedicated feature like [HONORIFIC], [FORMAL], or [POLITE].

A simple instance of recruitment is the French pronoun *vous*, which possesses both grammatical meaning (when directed at a plurality of addressees, (1a)) and social meaning (when directed, politely, at a singular addressee, (1b)). Thus, *vous* in (1a) contributes grammatical information about how many addressees are present, while *vous* in (1b) contributes social information about the speaker and addressee—namely, that they stand in a relationship requiring the expression of politeness. Throughout, I intend *polite* or *honorific* as interchangeable shorthands for the types of social meaning that involves greater social distance, whether it is based on respect, formality, or ritual.

(1) ***Avez*** ***vous*** *le livre?*
have.PRES.2**PL** 2**PL** the book
a. grammatical: 'Do <u>you (plural)</u> have the book?'
b. social: 'Do <u>you (singular + polite)</u> have the book?'

Singular *tu* has a different meaning profile from its plural counterpart. While it contributes grammatical information about how many addressees are present (one, in (2a)), it cannot be felicitously used to express politeness, indicated by # in (2b).

(2) ***As*** ***tu*** *le livre?*
have.PRES.2**SG** 2**SG** the book
a. grammatical: 'Do <u>you (singular)</u> have the book?'
b. #social: 'Do <u>you (singular + polite)</u> have the book?'

Pre-theoretically, there are two possibilities for how the plural is "re-used" for politeness. One possibility is that *social meanings are represented grammatically*, stemming from a dedicated grammatical feature like [HON]. In French, then, the polite pronoun and the second person plural pronoun are merely syncretic, or accidentally homophonous, with each other, as in (3). This is a core assumption of much traditional work on polite meaning, with scholars have assumed a dedicated grammatical feature in their representations of polite pronouns: [HON] (Simon 2003; Ackema & Neeleman 2018), [STATUS] (Macaulay 2015), or [FORMAL] (Portner, Pak & Zanuttini 2019).

(3) Syncretism/Accidental homophony

	SG	PL
1	*je*	*nous*
2	*tu*	*vous*
3	*il/illes*	*elles*
HON	*vous*	*vous*

This dissertation will pursue an alternative possibility that *social meanings are represented extra-grammatically*, with no dedicated features in the core grammar. The upshot is that the polite pronoun has no inherent form of its own, and in order to express polite meaning, the plural pronoun is somehow targeted for recruitment. The pronominal paradigm is as indicated in (4).

(4) Recruitment

	SG	PL
1	*je*	*nous*
2	*tu*	$\boxed{\textit{vous}}$ ↔ POLITE
3	*il/illes*	*elles*

Already, we have one good reason to favor (4): in (1), we see that using the polite pronoun obligatorily triggers corresponding plural verbal agreement, showing that the pronoun *is* plural. Thus, it is more parsimonious to assume (4b), a system with no specific exponents for polite pronouns or polite agreement, but that selectively recruits the plural form for its own exponence. Otherwise, we would have to assume syncretism/accidental homophony also exists in the verbal agreement system in a way that perfectly mirrors that of the pronominal system.

Recruitment shapes up to be a solid analytical possibility for several other reasons. Aside from parsimony, this dissertation will show that (4) is also favored by robust typological asymmetries in the shapes of polite pronouns. Recruiting the plural is extremely common typologically, being also attested in Wolaytta (Ethiopia, Omotic) (5). Recruits lead double lives: when deployed socially, they display a mismatch between the feature value that the speaker intends (singular) and the feature value that they were born with (plural). Throughout, recruitment results in these pronouns having both grammatical (a) and social (b) meanings.

(5) ***ʔinté*** *miiCC-**ídeta**.*
2**PL** laugh-PF.**2PL**
a. 'You (plural) laughed.'
b. 'You (singular + polite) laughed.' (Wakasa 2008: 1081)

Cross-linguistically, polite pronouns assume many more guises than plural. French and Wolaytta are languages which recruit the number opposition for honorification. Languages like Central Alaskan Yupik, however, recruit its person opposition, so that the polite pronoun takes the shape of third person (6).

(6) *ner'-**uq**=am.*
eat-IND.**3**SG=again
a. 'S/he is eating again.'
b. 'You (singular + polite) are eating again.' (Miyaoka 2012: 876)

Languages like Caddo (Caddoan, Arkansas) recruit person-neutral forms, a group of pronouns which encompass impersonals, reflexives, and anaphors. These frequently glossed as 'someone', 'people', or 'one' in typological descriptions. For a lack of a better term, I will fourth person as a cover term for these forms[1]. We see that the impersonal is recruited for politeness, as in (7).

(7) *dikat- **yi-** 'adih- hah?*
what- **IMP-** go- HAB
a. 'Where is one going?'
b. 'Where are you (singular + polite) going?' (Chafe 1990: 65, adapted)

Lastly, there are languages which recruit the inclusive for honorification. Consider Fehan Tetun

[1]In the literature on Inuit languages, fourth person is used to refer to morphemes that indicate non-coreference between one third person referent and another. The lack of co-reference is not what I intend. Rather, I adopt the term "fourth person" from the Athabaskan literature, where it describes person-neutral forms used in generic statements, indefinite referents, and/or absent referents. Thus, fourth person is not a subtype of third person, but a neutral person category of its own.

(Timor, Malayo-Polynesian), a language which makes singular/plural and clusivity distinctions. The inclusive plural pronoun *ita* is used for polite address, as in (8).

(8) *Ama Liu, ha'u k-mai k-amán* ***ita***...
father Liu 1SG 1SG-come 1SG-accompany **1INCL.PL**
a. 'I have come to accompany <u>us two</u>.'
b. 'I have come to accompany <u>you (singular + polite)</u>.' ' (van Klinken 1999: 113)

Even as recruitment creates systemic ambiguities across ϕ-featural categories, it is not unconstrained. (5)–(8) are part of screamingly robust typological asymmetries within the typology. While it is logically possible that polite pronouns take the shape of singular and first/second persons, it is unattested. There are no languages which recruit the singular for honorifying groups (plural is used in these cases), and none which recruit first person forms for honorifying second persons (third persons are used instead). The typological approach is crucial, as it provides us with both positive and negative evidence and ensures robust generalizations downstream. The pronominal typology to be presented is original empirical work, and was mainly informed by descriptive grammars, typological overviews, and anthropological studies; native speakers were consulted whenever possible. To assure the reader of the typological robustness of the displayed recruitments, each recruitment is supplemented with a table of languages exhibiting it, emphasizing wide geographical and genetic range.

This dissertation presents a (long) thought experiment: What if we buy into an analysis with recruitment (4) instead of one with dedicated features for social meaning as in (3)? Recruitment opens up several cans of worms, but that is what the rest of the dissertation will be used to address.

First of all, recruitment better be empirically motivated; Chapter 2 shows that it is very much so, presenting a typology of polite pronouns in n >220 languages which span multiple genera, multiple ϕ-featural values, and multiple grammatical domains, with stark recruitment asymmetries to be established besides. Well, the recruiting mechanism better be spelt out and be independently motivated by something else in the core grammar; it is the task of Chapters 3 and 4 which discuss semantic defaults in relation to politeness phenomena, showing that person-neutral or number-neutral forms are well-suited as avoidance mechanisms in the service of negative politeness. Well then, the final analysis better have a good grip on linguistic variation because polite pronouns are not uniformly exponed across all languages; Chapter 5 delivers an Optimality-Theoretic implementation which derives both attested and unattested sides of the asymmetry. The conclusion assesses the results of the thought experiment.

Chapter 2

The Journey Starts: Big Politeness

The present chapter presents the typology of ϕ-featural proxies of politeness from >220 genetically and geographically distinct languages, establishing the following empirical patterns:

- Polite pronouns are very frequently obtained by recruitment of pre-existing grammatical forms that are intrinsically not about politeness (plural, third person, impersonal person, inclusive).
- Recruitment displays robust typological asymmetries—only some feature values are recruited to the exclusion of others.
- When multiple recruitments occur, a politeness cline can be assembled—some recruits are more polite than others.

To help us make sense of the phenomenon of recruitment, I introduce some terminology.

Recruitment creates a relationship between *proxy* (the feature value doing the approximating) and *proxand* (the meaning being approximated to)[1]. Thus, in French, we say that the second person plural pronoun *vous* is the proxy, and singular polite address is the proxand. For terminological parity, we will interchangeably term recruitment as *approximation*[2].

I define *proxy* as some grammatical feature which serves as a stand-in for meanings that are not supplied directly by the core grammar (in this case, for politeness). Deploying a proxy often results in signature mismatches, as we have seen. Using *vous* for politeness creates a number mismatch: even though *vous* is grammatically plural, polite speakers use it as a notional singular.

The crux of this chapter is that, while honorifics assume a diverse range of guises, this diversity is neither random nor unconstrained. The generalization that emerges is that languages widely recruit plural number, third/fourth person persons, and inclusives for honorification. While the recruitment of singular number and first/second persons is logically possible, these are typologically unattested. At this stage, there is no *a priori* reason for why these grammatical feature values are specifically recruited for politeness, but these asymmetries pave the way towards providing an explanation in Chapter 3.

Preliminaries

Before the typology is laid out, I disclose further terminological disclaimers and make some delimitations regarding empirical scope. I define "politeness", and stress that this chapter is focused on the ϕ-featural reflexes of politeness. For those expressions of politeness which are not straightforwardly recruitment-based or ϕ-featurally based, such as allocutive agreement and differentiated speech registers, I will merely sketch them here, leaving them to be addressed fully in future work.

"Politeness": A disclaimer. What I interchangeably term *polite* or *honorific* forms are not meant to be interpreted as indications of true respect or reverence towards an individual. These labels merely serve as shorthands for the kind of distancing mechanisms involved in formal contexts,

[1] I adopt this terminology from terms like *divider–dividend* from mathematics, which Jaynes (1976) extends to *metaphor–metaphrand*.

[2] Recruitment also creates a continuous tension between the synchrony and diachrony of a language. When a plural is repeatedly used to mean singular, is it still a plural? Perhaps: politeness-driven repluralization is certainly highly frequent, where a new plural is innovated to fill the gap. It is also an open question as to whether this process is fully "online", so that the analysis of recruitment should be fully situated in the synchrony, or whether it has merely been conventionalized and be situated in the diachrony. I suspect it is somewhere in between; however, readers interested in the diachronic aspects of semantic reanalysis are referred to Eckhardt (2011) for an excellent overview.

or of marking an increase in social distance: *polite* or *honorific* forms carry an extra layer of social meaning, whereas their plain counterparts do not. How formality or deixis is defined is beyond the scope of the grammatical architecture itself.

Dramatically, there is evidence of person/number mismatches used towards individuals of inferior status. In pre-reform Swedish (Clyne et al 2006: 291), the 2PL pronoun *ni* was directed to those who were of lower social tiers, the opposite expected from "polite" uses of plural. (For expressing upward social distance, titles such as *Herr/Fru X* 'Mr./Mrs.X' were used.) Here, we see the determining factor for using the "polite" plural is distance from one's inner circle, not true reverence.

Different social groups deploy honorifics in different contexts: a wide variety of factors such as age, culture, social role, kinship, caste, social setting undeniably influence the deployment of the forms of interest. In most European languages, the honorific form is used towards individuals of higher social standing and/or non-intimates. In other languages, the use of honorifics may be additionally governed by factors such as age (e.g. in Korean, Acehnese), caste (e.g. in South India), discourse context (whether the conversation takes place in a casual, formal, or ceremonial setting), familiarity (e.g. in Polish), and kinship relations (e.g. in Aboriginal Australia). Hence, while speakers of French direct honorifics towards persons of higher social standing and non-intimates, speakers of Guugu Yimidhirr, for example, only do so towards fathers-in-law and brothers-in-law. Alternatively, honorifics may be directed towards situations instead of individuals: speakers of Warlpiri use certain honorifics only in ceremonial contexts (specifically, in initiation ceremonies, where boys are formally initiated into manhood). Moreover, the use of honorifics may be either reciprocal (e.g. in in-law avoidance registers) or non-reciprocal (e.g. in South India). In all cases, though, the absence of an expected honorific is deemed inappropriate or offensive.

The variables governing the *use conditions* of polite pronouns lie outside the scope of this work; what is crucial for our purposes is that, no matter the range of variation, the *morphosyntactic shapes* of polite pronouns are constrained by the proposal. The motivations behind using a distancing mechanism may vary widely, but *how* a distancing mechanism is chosen stays constant.

Allocutivity. Allocutive agreement is clause-final agreement markers which signal politeness exclusively geared towards addressees. In Japanese, *-masu* signals politeness towards the addressee, and has been analyzed by Miyagawa (2017) as agreement at C.

(1) *Otooto-wa* *ki-**mas**-u.*
younger.brother-TOP come-**ALLOC**-PRES
'My younger brother will come.' (Miyagawa, 2017: 20, adapted)

In Souletin Basque (Oyharçabal 1993), politeness is not the only social factor indexed by allocutive agreement; the sex of the addressee may be encoded via the same means. While *-ü-* encodes politeness towards an addressee of either sex, *-k-* is used for addressing a male colloquially and *-n-* a female colloquially. Allocutive agreement is also found in several other languages, including Jingpo (Myanmar, Tibeto-Burman, Zu 2013), Tamil (India/Sri Lanka, Dravidian, McFadden 2017), and Magahi (India, Indo-Aryan, Alok & Baker 2018).

Most allocutive markers cannot be (synchronically) derived via recruitment. However, some allocutive markers plainly recruit plural forms; I merely touch on this fact here. In Tamil, *-ŋgæ* is a marker of plurality in several areas, appearing in 2PL, 3PL verbal agreement, on nouns, and 2PL imperatives (Alok & Haddican 2022: 18-19, citing McFadden 2020). Notably, *-ŋgæ* is also an allocutive marker that is ambiguous between multiple familiar and singular respected (2).

(2) *Naan jaangiri vaang-in-een-**ŋgœ**.*
I jangri buy-PST-1SG.SBJ-PL
'I bought Jangri (spoken to multiple familiar addressees, or one respected addressee).'

Panjabi makes available a morphological distinction between singular vs. plural allocutive markers. Notably, it is also the plural allocutive marker that exhibits the same ambiguity (Kaur & Yamada 2022: 4).

(3) a. *Aman kitaab paRh reyaa* ***ii***
Aman book read PROG.M.SG ALLOC.SG
'Aman is reading a book (spoken to a familiar addressee).'
b. *Aman kitaab paRh reyaa* ***je***
Aman book read PROG.M.SG ALLOC.PL
'Aman is reading a book (spoken to multiple familiar addressees, or one respected addressee).' (Alok & Haddican 2022: 5)

Speech levels/Register. Lexical reflexes of politeness are probably universally present. One does not need to look to appreciate this kind of register-based variation. In English, the verb *to eat* has many socially conditioned variants even when we restrict ourselves to predicates which stay neutral about the speed or quantity of ingestion: *to partake, to dine* being formal variants whilst *to tuck into, to nosh on* are informal variants.

Looking further, Javanese (Indonesia/Malaysia, Malayo-Polynesian) distinguishes basic (*ngoko*), polite (*krama*), and super-polite (*krama inggil*) speech styles, with an intermediate style (*madya*). The forms of pronouns and certain predicates differ across speech styles. This is illustrated with variants of the sentence 'I see you', which are truth-conditionally equivalent but not socially so (Suharno 1982: 113, adapted):

(4)

'I	see	you.'	
aku	*wəroh*	*kowe.*	basic (*ngoko*)
kulo	*sumərəp*	*sampeyan.*	polite (*krama*)
daləm	*sumərəp*	*panjənənan.*	super-polite (*krama inggil*)

In the honorific register of Pohnpeian (Micronesia, Oceanic), certain predicates differ lexically depending on the social status of the addressee. This is shown for 'to eat' below (Keating & Duranti 2006: 151):

(5)

Commoner	*mwenge*
Low-status	*tungoal*
Chief	*koanoat*
Chieftess	*pwenieu*
2nd chief/chieftess	*sak*

Since registers of different formalities are not distinguished via ϕ-features, they are merely acknowledged here, and do not receive further attention in this dissertation.

Roadmap

Now that we have delimited the empirical scope of this chapter to polite pronouns, we are ready to move onto the full typology. §2.1 focuses on languages which feature one level of politeness, illustrating in detail the proxies of number, person, and clusivity. §2.2 then moves on to languages

which feature multiple levels of politeness, showing that the same recruitment asymmetries pertain, and additionally establishing clines in politeness—some recruits are better-suited for higher levels of politeness. §2.3 shows that recruitment displays robust typological asymmetries, so that certain recruitment profiles are unattested, even if they are logically possible.

Hence, politeness pressures, whatever they are, establish proxy-proxand relationships; the grammar does not supply dedicated exponents for polite pronouns. In this state of affairs, the only recourse available to the grammar is approximation. In what follows, the additional layer of social meaning—politeness—will be indicated in the translation line with (POLITE). The absence of (POLITE) indicates that a pronoun does not carry polite meaning.

2.1 Single-tier politeness systems

This section offers a typologically extensive tour of languages featuring one level of politeness, making the binary distinction between polite *vs.* non-polite. §1.1–§1.3 detail the honorific proxies of plural, third/fourth persons, and inclusives, for the proxand of singular polite address. §1.4 shows that similar recruitments pertain for other proxands: singular polite reference, singular polite/humble self-reference, and plural polite address. §2.1.5 establishes the typological asymmetry itself, illustrating logically possible yet unattested politeness systems. §2.1.6 illustrates single-tier systems exhibiting multiple proxies in free variation.

2.1.1 Honorific uses of number

Recall (1), repeated below as (6), showing that plural number is recruited for honorification in French. For polite address towards a singular addressee, the 2PL pronoun *vous* is used, creating a mismatch between literal meaning (plural) and conveyed meaning (singular and honorific).

(6) French honorific plural
Avez ***vous*** *le livre?*
have.PRES.**2PL** **2PL** the book
'Do you (POLITE) have the book?'
Lit.: 'Do <u>you (plural)</u> have the book?'

This pattern is well-attested for several other European languages, where 2PL pronouns can be used to respectfully address one person. In some cases, the respectful form is capitalized in the orthography[3]. Such pronouns include Lithuanian *jūs*; Swedish *Ni*; Russian *Vy*; Slovenian *Vi*, Sorbian *wy*; Czech, Slovak, Bosnian *vy*; Serbian *vi*; Italian (southern dialects) *voi*; Belarusian *vei*; Ukrainian, Macedonian, Bulgarian *vie*; and Finnish *te*.

However, honorific plural is not restricted to European languages. It is in fact a cross-linguistic trend, being the most frequent honorification strategy across the world's languages, found in geographically and genetically unrelated languages. The singular pronoun is used for plain address, while the plural pronoun is used to add an additional layer of polite meaning. If verbal agreement is morphologically distinguishes between number, we see that it tracks the number features of the pronoun. We see precisely this pattern in Khimt'anga/Xamtanga (Ethiopia, Cushitic) (7)[4], Tamil (India/Sri Lanka, Dravidian) (8), Wolaytta (Ethiopia, Omotic) (9) and in Persian (Iran/Israel, Indo-European) (10)[5].

[3]Sometimes, both familiar and polite forms are capitalized, such as German familiar *Du* and polite *Sie*.
[4]In this example, EP is the abbreviation for an epenthetic vowel.
[5]Many thanks to Omri Doron with help in glossing Persian examples.

(7) Khimt'anga/Xamtanga honorific plural

a. ***kɨt***-∅ *tə-* *tɨr -u.*
2SG.S-NOM come- 2- PERF
'You came.'

b. ***kɨtɨn***-∅ *tə-* *tɨr- i-* ***n*** *-u.*
2**PL**.S-NOM come- 2- EP- **PL** PERF
'You (POLITE) came.'
Lit.: 'You (plural) came.' (Belay 2015: 158-9)

(8) Tamil honorific plural

niin-kaḷ *eppootu va-nt-**iirkaḷ**.*
2-**PL** when come-PAST-2**PL**
'When did you (POLITE) come?'
Lit.: 'When did you (plural) come?' (Lehmann 1993: 100-1)

(9) Wolaytta honorific plural

ʔinté *miiCC-**ídeta**.*
2**PL** laugh-PF.**2PL**
'You (POLITE) laughed.'
Lit.: 'You (plural) laughed.' (Wakasa 2008: 1081)

(10) Persian honorific plural

shomā *be dust-**etān*** *goft-**id**?*
2**PL** to friend-2**PL**.GEN tell.PST-2**PL**
'Did you (POLITE) tell your friend?
Lit.: 'Did you (plural) tell your friend?' (Yousef 2018: 60-5, adapted)

In Warlpiri (Australia, Pama-Nyungan), different-sex siblings and co-initiates (those who were shared an initiation ceremony with oneself, translated here as 'agemate-brother) are two categories of persons warranting the use of the respect register. In (1), we see that addressing a singular coinitiate is done so with the plural.

(11) Warlpiri honorific plural

*ngana-ngku-**nyarra*** *ngarrurnu yarlpurru-pardu?*
who-ERG-**2PL**.OBJ told coinitiate-DEF
'Who told you (POLITE), (my) agemate-brother?'
Lit.: 'Who told you (plural)?' (Laughren 2001: 210-1, adapted)

Table 2.1 shows that honorific plural is extremely well-represented, robustly present in a geographically and genetically diverse group of languages. Thus, the honorific use of plural is by no means an isolated accident, but a highly robust typological trend.

The diacritics on several languages in Table 2.1 mark number systems aside from singular-plural: number systems can be more articulated, with multiple nonsingulars such as dual (for 'two') and paucal (for 'a few'). * indicates that the language has a singular-dual-plural system, ** a singular-dual-paucal-plural system, *** a singular-dual-paucal-greater paucal-plural system, + a minimal-augmented system, and ++ a minimal-unit augmented-augmented system[6].

Recent work in number theory captures the diversity of number systems with different featural representations for each number system (Hale 1973; Silverstein 1976; Noyer 1992; Harbour

[6]Minimal-augmented pronominal systems include an additional first person pronoun, a 'minimal' pronoun. This pronoun exclusively refers to the speaker-hearer dyad. An 'augmented' pronoun simply adds a number of others to this dyad. This is exemplified with the independent pronouns of Muna below (van den Berg 2013: 51):

Family	Language(s)
Athabaskan-Eyak-Tlingit	NE. Athabaskan
Austroasiatic	Jahai*, Khasi, Temiar*
Bantu	Bena, Bemba, Chichewa, Fwe, Makonde, Northern Sotho (Pedi), Nsenga, Silozi, Tsotsitaal, Zulu
Baltic	Lithuanian*
Benue-Congo	Yoruba
Bunaban	Bunaba, Gooniyandi
Caddoan	Caddo*
Cariban	Galibi Carib^{+}, Wai Wai^{+}
Celtic	Scottish Gaelic, Welsh
Chadic	Bachamar
Cushitic	Hdi^{+}, Mina^{+}, Kambaata, Khimt'anga, Wandala
Dravidian	Malayalam, Kannada, Tamil, Telugu
Germanic	Swedish
Graeco-Phrygian	Modern Greek
Gunwinyguan	Warray
Indo-Aryan	Assamese, Eastern Awadhi, Bangla, Bhojpuri, Gujarati, Hindko, Hindi-Urdu, Kannada, Kashmiri, Maithili, Marathi, Marwari, Oriya, Panjabi, Saraiki
Indo-European	Armenian (Western)
Indo-Iranian	Persian
isolates	Ainu^{+}, Burushaski (Eastern)
Kartvelian	Georgian
Khoisan	Nama*
Kwa Volta-Congo	Tafi
Malayo-Polynesian	Duri, Mori Bawah, Pangasinan^{+}, Tagalog, Tukang Besi, Waikyon Taba
Mangarayi-Marran	Mangarayi*
Micronesian	Ponapean*
Nakh-Daghestanian	Lezgian
North Volta-Congo	Dii^{+}, Gbaya-Manza-Ngbaka, Koromfé, Kolbila, Mayogo, Sango
Oceanic	Boumaa Fijian**, Dii^{+}, Efate (South)*, Gbaya-Manza-Ngbaka, Iduna, Imere*, Kaulong*, Kobon*, Kolbila, Koromfé, Mayogo, Sango, Standard Fijian**, Sursurunga***, Tamambo, Tigak*, Tinrin*, Usan
Omotic	Haro, Wolaytta
Otomanguean	Coatzospan Mixtec, Silacayoapan Mixtec
Pama-Nyungan	Djambarrpuyngu^{++}, Djaru*, Guugu Yimidhirr*, Mparntwe Arrernte*, Nyangumarda*, Warlpiri*, Martuthunira^{++}, Kuku-Yalanji*, Walpiri*, Wik-Ngathana*
Romance	French
Siouan	Assiniboine
Sino-Tibetan	Camling*, Cogtse Gyarong*, Dhimal*
Slavic	Belarusian, Bulgarian, Czech, Latvian, Macedonian, Russian, Serbian-Croatian-Bosnian, Sorbian*, Slovak, Slovenian*, Serbian
Turkic	Turkish, Tuvan
Uralic	Finnish

Table 2.1: Languages with honorific plural

2014). If this line of thinking is on the right track, then Table 2.1 also displays the *featural* diversity of honorific plural: the "plurals" that are put to honorific uses may not stem from identical formal origins. For instance, the Caddo honorific plural is recruited from a singular-dual-plural system, while the Oroha honorific plural is recruited from a singular-dual-paucal-plural system. (There are also languages utilizing their duals for honorification; interested readers are referred to Wang 2023: §5.) Since these plurals are featurally distinct, this suggests that the phenomenon of honorific plural is not dependent on how a particular language's plural is formally derived.

Other than being a numerically robust pattern, honorific plural is also *grammatically* robust, crosscutting several grammatical domains within a language. Thus, in any environment that singular and plural numbers are morphologically distinguished, plural may approximate singular polite address. Above, this has been illustrated with the shape of full pronouns and verbal agreement, but also extends to any environment where the relevant number distinctions are made. Consider Tamil. We have already seen honorific plural in full pronouns and the corresponding verbal agreement in (8). However, the same distinction is made in imperatives (12), as well as in prohibitives (13).

(12) a. *Keeḷ-∅*
Ask-2
'(You) Ask!'

b. *Keeḷ-**uṅkaḷ***
Walk-**PL**
'Please ask!'
Lit.: '<u>You (plural)</u> ask!'

(13) a. *Keeṭk-aat-ee*
ask-NEG-EUPHONIC
'Don't ask!'

b. *Keeṭk-aat-**iirkaḷ***
ask-NEG-**2PL**
'Please don't ask!'
Lit.: '<u>You (plural)</u> don't ask!'

	Minimal		Augmented	
1EXCL	*inodi*	'I'	*insaidi*	'I and others'
1INCL	*intaidi*	'You and I'	*indaidi-imu*	'You and I and others'
2	*ihintu*	'You'	*ihintu-umu*	'You and others'
3	*anoa*	'Him/her'	*andoa*	'Them'

Thus, minimal-augmented systems are characterized by an inconsistent referential cardinality within each number. Minimal pronouns denote dual in the 1INCL person (denoting the speaker-hearer dyad), singular otherwise. Augmented pronouns have a referential cardinality of >2 in the 1INCL person, >1 otherwise.

Unit augmented is an intermediate number category between minimal and augmented, whose cardinality is 1 more than that of the minimal category. Thus, unit-augmented pronouns denote a trial if they contain the speaker-hearer dyad (and one other), but dual otherwise. This is exemplified by the independent pronouns of Wanyjirra (Australia, Pama-Nyungan, Senge 2015: 214):

	Minimal		Unit augm.		Augmented	
1EXCL	*ngayu*	'I'	*ngaliyarra*	'I and other'	*nganimba, ngadiba*	'I and others'
1INCL	*ngali*	'You and I'	*ngaliwula*	'You and I and other'	*ngaliwa*	'You and I and others'
2	*nyundu*	'You'	*nyunbula*	'You and other'	*nyurrara*	'You and others'
3	*nyandu*	'Him/her'	*nyanbula*	'Him/her and other'	*nyarralu*	'Him/her and others'

2.1.2 Honorific uses of person

Another grammatical category recruited for honorification is person. Italian recruits third person, creating a person mismatch between literal meaning (third person) and conveyed meaning (second person and honorific), as in (14) below[7]

(14) Italian honorific third person
Signor Alessandro, ***è/*sei*** *contento?*
mister A **3**SG/**2**SG.COP happy.MASC
'Mr. Alessandro, are you (POLITE) happy?'
Lit.: 'Sir Alessandro, is <u>he</u> happy?'

The same pattern is also found in Central Alaskan Yupik (Alaska, Yupik) (15).

(15) Central Alaskan Yupik honorific third person
*ner'-**uq**=am!* *(=nìɣɣùqqam)*
eat-IND.**3**SG=again
'You (POLITE) are eating again!'
Lit.: '<u>S/he</u> is eating again.' (Miyaoka 2012: 876)

In Warlpiri, third person forms may be used for addressing individuals warranting avoidance: in this case, towards older men who were present at one's initiation ceremony. Thus, (16a) features the second person affix *-npa* and is used for address in the regular register, but (16b) features the unmarked third person affix and is used for address in the avoidance register. (In the avoidance register, verbs also undergo semantic bleaching: thus *nyina* 'to live, stay, sit' is changed into a general spatial verb *karrka*, which does not lexically specify the nature of how one is being located. For further discussion of such instances of lexical bleaching, see Wang 2023: §4.4.1.)

(16) Warlpiri honorific third person
a. *Nyampu-rla-juku mayi ka-**npa*** *nyina?*
here-LOC-still Q PRES-**2**.SUBJ live
'Are you still living here?'
b. *Nyampu-rla-juku mayi ka-∅* *karrka?*
here-LOC-still Q PRES-**3**.SUBJ locate
'Are you (POLITE) still living here?'
Lit.: 'Is <u>s/he</u> still living here?' (Laughren 2001: 214-5)

The number of languages which recruit third person are much fewer than those which recruit honorific plural. Nonetheless, the relevant languages do not constitute a geographically or genetically homogeneous group, as Table 2.2 shows.

The last type of honorification strategy to be presented is that of honorific fourth person. For my purposes, fourth person forms are pronoun forms which suffer from an absolute poverty of person features, not making reference to any discourse participant, unlike the first, second, or third persons, so that they have been variably described as impersonal, indefinite, and generic. Here, I expand the category here to any person-agnostic form, including anaphors like 'self'.

Ainu (Hokkaido, isolate) features an impersonal pronoun *a(n)*, used for indeterminate ref-

[7]This example is one of polite singular address towards a male. When addressing a female, gender on the adjective becomes feminine: *Signora Alessandra, e contenta?* This suggests that the polite pronoun *Lei* is not itself specified for gender, despite its resemblance to the current third person singular feminine pronoun.

erents (17a). We see that it is this pronoun which is recruited for honorification (17b).

(17) Ainu honorific impersonal

a. ***An*** *omap pon menoko ne ruwene.*
IMPERS love be.young woman COP ASSERT
'She is a young woman whom *people* like.'

b. ***An*** *nu no.oka ...*
IMPERS ask IMPF
'As you (POLITE) are asking...'
Lit.: 'As <u>someone</u> is asking...' (adapted from Refsing 1986: 219, 222)

In Caddo (Caddoan, Arkansas), the prefix *yi-*, termed a "defocusing prefix", may be used for an impersonal reading. However, it is also used in the respect register, reserved for addressing one's singular in-law. In the regular register, the affix specified for second person is used (18a), but in the respect register, the impersonal *yi-* is used in place of the ϕ-featurally specified agreement affix (18b).

(18) Caddo honorific impersonal

a. *dikat-* ***yah'-*** *'adih- hah? (=dikadàadihah)*
what- **2-** go- HAB
'Where are you going?'

b. *dikat-* ***yi-*** *'adih- hah? (=dikadiidihah)*
what- **IMPERS-** go- HAB
'Where are you (POLITE) going?'
Lit.: Where is <u>one</u> going?' (Chafe 1990: 65, adapted)

In Angolar Creole Portuguese, fourth person also has this use, as (19a) shows the fourth person pronoun *a* being used for polite address. In (19b), *a* co-occurs with the respectful address term su^n 'mister' to "convey a more intense sense of respect towards the addressee". In both cases, *a* is used instead of 2SG *bo*.

(19) Angolar Creole Portuguese honorific impersonal

a. ***a*** *Ta ku e kikie bende?*
IMPERS be with 3SG fish sell
'Do you (POLITE) have fish to sell?'
Lit.: 'Does <u>one</u> have fish to sell?' (Lorenzino 1998: 154 citing Maurer 1995: 62)

b. *Antigamente* su^n ***a*** *ka kono izakenci,* ***a*** *ka ba kw'e*
Long.time.ago mister **IMPERS** HAB take izaquente **IMPERS** HAB go with-3SG
awa, a ka laba.
water **IMPERS** HAB wash
'In the old days you (POLITE) took the izaquentes and washed them with water.'
Lit.: 'In the old days <u>one</u>...' (Lorenzino 1998: 145)

In Navajo (Athabaskan-Eyak-Tlingit, Arizona), fourth person is used for unspecified or indefinite subjects, or for politely referring to a discourse participant (Faltz 1997: 32). This use of fourth person also pertains in the related language Wailaki (California USA, Athabaskan-Eyak-Tlingit), where *ch'i-* can be used for either purpose (Begay 2017: 174).

We see the proxy of fourth person in imperative constructions as well. In Kelabit (Malayo-Polynesian, Borneo), the fourth person pronoun *narih* is the pro-form used in generic and prescriptive statements. It is *narih* which expones the grammatical subject in polite imperatives

(20).

(20) *Laak-en* ***narih*** *si'it nubaq na'an.*
cook-UV.IRR IMPERS little rice later
'Could you cook a little rice later.' (Hemmings 2016: 137; p.c.)

Lastly, closer to home, the pronoun *on* in French, variably described as a impersonal, generic, or indefinite pronoun, are found in polite prescriptions and prohibitions. Both (21a-b) can be directed towards a specific addressee, regardless of whether the ϕ-featurally specified pronoun is used or whether the impersonal is used. However, (21a) is better directed towards a child, while (21b) is directed towards e.g. the President.

(21) French polite fourth person

a. ***Tu*** *ne dois pas marcher sur l'herbe.*
2SG NEG must.2SG NEG walk on the.grass
'You should not walk on the grass.'

b. ***On*** *ne doit pas marcher sur l'herbe.*
IMPERS NEG must.3SG NEG walk on the.grass
'You (*Lit.:* <u>One</u>) should not walk on the grass.' (Adèle Hénot-Mortier, p.c.)

In Yiddish, the impersonal *men* can also be used for polite address (alongside 2PL *ir*), albeit as part of a formal register.

(22) *Ven ot di yediyes ken* ***me*** *nisht krign, ...*
if DEM the notices can **IMPERS** not obtain
'If you (formal) can't get that information...' (Adam Albright, p.c.)

Table 2.3 summarizes the languages with honorific fourth person found in my sample. It is again notable that these languages are geographically and genetically heterogeneous.

2.1.2.1 Combinations of values

So far, we have seen that the feature values of plural, third person, or fourth person is recruited for singular polite address. However, languages can also use a combination of feature values for politeness.

The third person plural

Here, the two strategies of honorific plural and honorific third person are combined. The result is that 3PL pronouns are used for singular honorific address. A well-known instance of this is German, where its 3PL pronoun *Sie* is recruited for singular honorific address (23). This obligatorily triggers 3PL verbal agreement, underlining the fact that this is true recruitment, not accidental homophony. Furthermore, while 3PL *Sie* is phonologically identical to 3SG.FEM *sie*, verbal agreement shows us that it is the former being recruited. The mismatch created is twofold: in terms of both person and number. so that the literal meaning (third person plural) diverges from the conveyed meaning (second person singular, and polite).

(23) German honorific third person plural
Sind/*Bist ***Sie*** *müde?*
3PL/2SG.COP **3PL** tired
'Are you (POLITE) tired?'

Family	Language(s)
Austroasiatic	Jahai, Semelai
Bunaban	Bunaba
Tangkic	Lardil
Pama-Nyungan	Warlpiri
Inuit	N. S. Iñupiaq
Yupik	Central Alaskan Yupik
isolate	E./W. Basque, Ktunaxa
Romance	Italian
Uto-Aztecan	Tlaxcala-Puebla-C. Nahuatl
Sino-Tibetan	Eastern Kayah Li

Table 2.2: Languages with honorific third person

Family	Language(s)
Creole	Angolar Creole Portuguese
Cushitic	Kambaata
Khoisan	Khwe
Athabaskan-Eyak	Navajo, Wailaki, W. Apache
Caddoan	Caddo
Germanic	Irish English, Yiddish
Otomanguean	Silacayoapan Mixtec
Indo-Aryan	Assamese, Bangla, Bhojpuri, Gujarati, Hindi-Urdu, Kannada, Kashmiri, Maithili, Marwari, Oriya
Indo-European	English
isolate	Ainu
Romance	French
Malayo-Polynesian	Acehnese, Kelabit, Lun Bawang
Mongolic	Mongolian
Quechuan	Quechua (Bolivian; Imbabura)
Sino-Tibetan	Tamang

Table 2.3: Languages with honorific fourth person

Family	Language(s)
Bantu	Bemba
Germanic	German, Danish, Norwegian
Slavic	Slovenian (dialectic)
Oceanic	Ponapean
Austronesian	Tagalog

Table 2.4: Languages with honorific third person plural

Lit.: 'Are <u>they</u> tired?'

The same recruitment strategy is found in Norwegian, where the 3PL pronoun, *De*, is recruited for honorific address.

(24) Norwegian honorific third person plural
Er ***De*** *syk?*
COP **3PL** sick.SG
'Are you (POLITE) sick?'
Lit.: 'Are <u>they</u> sick?' (Nathan Young, p.c.)

Table 2.4 shows the spread of languages which use third person plural for singular honorific address.

The third person reflexive

So far, the fourth person forms we have seen recruited for politeness have been person-neutral. However, some languages have person-specified reflexives; one such example is Hungarian. To form reflexive pronouns, the stem *mag-* is combined with person/number suffixes from the possessive series, as in (25).

(25) Hungarian reflexive pronouns (nominative) (Rounds 2009: 128)

	SG	PL
1	*magam*	*magunk*
2	*magad*	*magatok*
3	*maga*	*maguk*

Of these, third person reflexives are the ones recruited for politeness or social distance: 3SG.REFLX *maga* for singular address, and 3PL.REFLX *maguk* for plural address. They are "used to address people you are not as familiar with as well as strangers with whom you wish to maintain a reserved but not too formal relation." (26a) illustrates the reflexive use of *maga*, while (26b) its polite use.

(26) Hungarian polite third person reflexive (Rounds 2009: 126, 129, adapted)

a. *Illona nem szeret* ***mag****-áról* *beszél-ni.*
Illona NEG like **3SG.REFLX**-about speak-INF
'Ilona doesn't like to speak about herself.'

b. ***(Maga)*** *milyen* *könyvet keres?*
3SG.REFLX what.kind.of book search

'What kind of book are you (POLITE) looking for?'

2.1.3 Honorific uses of inclusive

Recall Fehan Tetun (Timor, Malayo-Polynesian), a language which makes singular/plural and clusivity distinctions. The 1INCL.PL *ita* is recruited for addressee honorifics. This was (8), repeated as (27) below.

(27) Fehan Tetun honorific inclusive (van Klinken 1999: 112-3)
Ama Liu, ha'u k-mai k-amán ***ita****...*
father Liu 1SG 1SG-come 1SG-accompany **1INCL.PL**
'Father Liu, I have come to accompany you (POLITE)... '
Literally.: 'I have come to accompany me and you...'

In Guarijio (Mexico, Uto-Aztecan), the inclusive is also recruited for honorification. In the following minimal pair, (28a) with the second person *múú* does not carry polite meaning, while (28b) with the inclusive *remé* does.

(28) Guarijio honorific inclusive
a. *Muké-ru* ***muú*** *ku'írabi?*
carry-PAST.PASS **2** short.while
'Did you carry it on your shoulder a while?'
b. *Muké-ru* ***remé*** *ku'írabi?*
carry-PAST.PASS **1INCL** short.while
'Did you (POLITE) carry it on your (POLITE) shoulder a while?'
Lit.: "Did we carry it on our shoulder a while?' (Miller 1980: 197)

In Yagua (Peru, Peba-Yagua), polite imperatives employ the first person plural inclusive clitic *vuurya* even if they are directed towards a second person, as in (29).

(29) ***Vuurya****-a-maa jaachiy-pííy-yaa.*
1INCL-IRR-PERF heart-VERBLZ-DIST
'You (POLITE) had better think (about something).'
Lit.: 'We had better think.' (Payne 1985: 65)

In Cebuano (Philippines, Malayo-Polynesian), polite imperatives have a similar makeup. Usage of a second person is not polite, while usage of the inclusive *ato* is[8].

(30) a. *basta unsa=y imo-ng(a) plano, ∅ padayon-a.*
PART what=INDEF 2SG.POSS-LINK plan 2SG continue-PV.IMPERF
'Whatever your plan is, (you) go ahead (with it).'
b. *basta unsa=y imo-ng plano,* ***ato****-ng(a) padayon-on.*
PART what=INDEF 2SG.POSS-LINK plan **1INCL**-LINK continue-PV
'Whatever your plan is, you (POLITE) go ahead (with it).'
Lit.: 'Whatever your plan is, we'll go ahead (with it).' (Tanangkingsing 2009: 266)

Honorific uses of the clusivity distinction are well-supported typologically, as Table 2.5

[8]In this example, PV is the abbreviation for Patient Voice. Furthermore, in both the non-polite (a) and polite (b) examples, note that the possessive phrase remains second person singular: *imo-ng plano* 'your plan'

shows.

Family	Language(s)
isolate	Ainu
Indo-Aryan	Gujarati
	Marathi
	Marwari
Malayo-Polynesian	Äiwoo
	Bajau
	Batak Toba
	Buginese
	Cebuano
	Duri
	Muna
	Konjo
	Makasar
	Niuafo'ou
	Pukapukan
	Sabah Malay
	Sawu
	Tehun (Fehan)
	Tikopia
	Tukang Besi
	Wuvulu
North Halmahera	Galela
Oceanic	Kilivila, Tongan
Peba-Yagua	Yagua
Quechuan	Cuzco Quechua
Sino-Tibetan	Limbu

Table 2.5: Languages with honorific clusivity

We do not find languages recruiting the exclusive for honorification. The only potential candidate for this is Silacayoapan Mixtec (Mexico, Otomanguean), where Shields (1988: 407) reports that both plural and exclusive can be used for respect, explicitly stating that the inclusive does not have this use. However, throughout the source, only instances of honorific plural with 2PL *ndóhó* were found, without instances of honorific exclusive with 1EXCL.PL *ndehe*. Due to this lack of direct evidence, we exclude it from the sample.

2.1.4 Distinct proxands of politeness

So far, we have only covered one proxand: politeness directed towards a second person (*singular polite address*), where various proxies substitute for the 2SG form, uncovering the recruitment asymmetries summarized in Table 2.6.

Recruitment asymmetries also pertain for other proxands: politeness directed towards third persons (*polite reference*) and first persons (polite self-reference, or, more precisely, *humbling*). We will cover these cases now.

For a single recruit, languages differ on whether it may be used for self-address, address, and/or reference[9]. Tinrin allows honorific plural for singular address, but not honorific reference

[9]If a language uses honorific plural for e.g. both address and reference, it is only listed—and counted—once in the typology.

Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	125	27	12	7	29
Language	–	–	–	French	Fehan Tetun	Italian	German	Ainu
				vous	*ita*	*Lei*	*Sie*	*an*

Table 2.6: Recruitment asymmetries: Singular polite address

(Osumi 1995). Indian languages such as Kashmiri (Koul & Wali 2006: 51), Kannada (Schiffman 1983: 38), Tamil (Lehmann 1993: 100-1), and Malayalam (Asher & Kumari 1997) allow for both. Koromfé (Burkina Faso, Volta-Congo) recruits the plural for addressing, or referring to, individuals such as village chiefs or God (Rennison 1997: 246).

As far as the data shows, there is no implicational hierarchy to be established, only that politeness directed towards second persons is the most frequent. This is probably due to its relatively high utility: referents are not necessarily physically present in the same discourse, while addressees definitely are.

It is reasonable to think that a language will possess grammatical means of expressing politeness towards non-second persons if it already has those for expressing politeness towards second persons. However, San Carlos Western Apache (Arizona, Athabaskan) provides us with a counterexample: the language recruits fourth person for polite reference, but not for polite address. In those cases, the lack of politeness directed towards second person may be indicative of an extreme taboo: one is not allowed to address a tabooed social relation at all. Since address is already impossible, so is polite address.

2.1.4.1 Politeness directed to a third person

When a speaker is merely referring to a third person who may not even be present in the current conversation, one might wonder if similar politeness considerations even pertain.

I propose that yes—honorific reference should be derived exactly parallel to honorific address. Once a tabooed relation is established, it is equally prohibitive to refer to that relation as it is to address them. Both linguistic and anthropological work highlight this feature of taboo, particularly those of in-law taboos. Rushforth (1981:35-6) notes the following rules of in-law avoidance in the Northeast Athabaskan-speaking Bear Lake Dene community. At all times, one is to avoid unnecessary conversation with an in-law. If conversation is necessitated, it should be done indirectly, through an appropriate proxy. If an appropriate proxy cannot be found, only then can one speak directly to an in-law, but only in the affinal speech style, characterized by use of the honorific plural or honorific fourth person. Previous studies about the Dene ethnographic group in general (Helm 1961; Savishinsky 1970) are concordant regarding this practice. In Lardil (Norvin Richards, p.c.), the third person pronoun is used for honorific reference towards a tabooed relation, but the taboo extends far beyond speech: tabooed relations cannot be looked at, pointed at, or touched.

Similarly wide-ranging taboos can be found in many other language communities. In Guugu Yimidhirr, brothers-in-law "not only use the respectful vocabulary; they sit far apart, orient their bodies so as not to face one another, and avoid direct eye contact" (Haviland 1979: 170). In Warlpiri, tabooed relations are identified as in-laws, co-initiates, and opposite-sex relations (specifically opposite-sex siblings). These taboos are arbitrated by honorific plural, regardless of whether the recipient is an addressee or a referent. Laughren (1996: 192) notes that speakers "use the plural pronoun *nyurrurla* and not the singular *nyuntu(lu)* to address or refer to a different

Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	0	0	0	28	8
Language	–	–	–	–	–	–	Wolaytta	W. Apache
							ʔetaa	*ko-*

Table 2.7: Recruitment asymmetries: Singular polite reference

sex sibling."

Hence, the decision to mark social deixis between oneself and a non-intimate does not depend on whether a non-intimate in question is in direct earshot or not. Here, we will consider the various recruits which are substituting for the 3SG form. We see that the proxies for honorific reference exhibit a similar, but sharper, asymmetry. This is because, for polite reference, third person is not a recruit or proxy anymore, exhibiting no mismatch. The pool of recruits is much more restricted, as Table 2.7 shows: for singular polite reference, recruits are limited to the smaller pool containing the plural and fourth person.

Polite reference via plural

We now see that much of the data on honorific plural above showed an incomplete picture: while used for singular polite address, plural is also used for singular polite reference. In this case, the form of the polite pronoun is that of 3PL.

Recall from (1) that Warlpiri recruits plural for singular polite address towards coinitiates, repeated below as (31a). The same recruitment is possible for singular polite reference to coinitiates, shown in (31b), or to in-laws, as in (31c), where the speaker asks after his son-in-law with the plural. Note also his explicit avoidance to the "son-in-law" relation, which is only alluded to by pointing out that his addressee is his son-in-law's mother.

(31) Warlpiri honorific plural

a. *ngana-ngku-**nyarra** ngarrurnu yarlpurru-pardu?*
who-ERG-**2PL**.OBJ told coinitiate-DEF
'Who told you (POLITE), (my) agemate-brother?'

b. *Yuka-ya, kaji-ka-ngku-**lu** yarlpurru-pardu-rlu nya-nyi.*
enter-IMP COMP-PRES-2SG.OBJ-**PL**.SUBJ coinitiate-DEF-ERG see-PRES
'Get in, or (my) agemate-brother (POLITE) might see you.' (Laughren 2001: 210-1, adapted)

c. *Nyarrpara waja-**lu** ngarrijarri-ja nyuntu-npa japun-warnu-ju?*
where.NOM P-**PL.SUBJ** locate-PAST you-2.SUBJ mother-ASSOC-TOP
"Where is it he (POLITE) went? You're his mother."
Lit.: "Where did you say he was located that you are the mother of?" (Laughren 1996: 213)

The prevalence of honorific plural across address and reference also pertains for several other languages already mentioned. The plural that Wolaytta uses for honorific address (32a), repeated from (9), is also used for honorific reference (32b).

(32) Wolaytta honorific plural

a. ***ʔinté** miiCC-**ídeta**.*
2**PL** laugh-PF.**2PL**

'You (POLITE) laughed.'

b. *hagéé* ***7etaa****-g-áá.*
this 3**PL**-NOMLZ-ABS
'This is his (POLITE).' (Wakasa 2008: 1081)

Recall from (10) that in Persian, the plural is used for singular polite address, repeated below as (33a). However, it can also be recruited for polite reference (33b)[10]. In (33b), we see both polite address via 2PL *shomā* in the first conjunct, and polite reference via 3PL *ishān* in the second conjunct.

(33) Persian honorific plural

a. ***shomā*** *be dust-****etān*** *goft-****id****?*
2**PL** to friend-2**PL**.GEN tell.PST-2**PL**
'Did you (POLITE) tell your friend?

b. *pedar-e* ***shomā*** *ānjā bud-and, va* ***ishān*** *ham bā-man sohbatkard-and.*
father-POSS 2**PL** there COP-INF and 3**PL** also to-1SG speak-INF
'Your (POLITE) father was there, and he (POLITE) too talked to me.' (Yousef 2018: 60-5, adapted)

(34) Assiniboine honorific plural

a. *mišíc'e* *o'ínažį ektá* ∅*-yį-kta* ∅*-káya.*
my.brother.in.law town to A3-go-POT A3-say
'My brother-in-law said he is going to town.'

b. *mik^{h}ʉ* *o'ínažį* ∅*-yá-****pi****-kta* ∅*-káya-****pi****.*
my.mother.in.law town A3-go-**PL**-POT A3-say-**PL**
'My mother-in-law said she (POLITE) is going to town.' (Cumberland 2005: 146-7)

c. *Mi?ús,* *t^{h}anó né* *awámìciyàka* ***pó****...* *síca-****pi!***
my.buttock meat DET look.for.me **PL**.IMP be.bad-**PL**
'My Buttocks, watch over this meat for me... You (POLITE) are bad!' (Cumberland 2005: 148, 434-7, adapted)

Polite reference via fourth person

The impersonal affix that Caddo uses for singular polite address, shown above in (98a), is also used for polite reference. In the regular register, the affix specified for third person is used (35a), but in the respect register, the impersonal *yi-* is used instead (35b).

(35) Caddo honorific impersonal

a. *dikat-* ∅*- 'adih- hah? (=dikadàadihah)*
what- **3**- go- HAB
'Where is he/she going?'

b. *dikat-* ***yi-*** *'adih- hah? (=dikadiidihah)*
what- **IMPERS**- go- HAB
'Where is he/she (POLITE) going?' (Chafe 1990: 65, adapted)

In San Carlos Western Apache (Arizona, Athabaskan), the fourth person personal prefix, *ki-* ∼ *ko-*, is ambiguous between an indefinite and a referent honorific (de Reuse 2006: 79). There is no use of *ki-* ∼ *ko-* reported for addressee honorification.

[10]Many thanks to Omri Doron with help in glossing.

(36) ***ko**-god*
IMPERS-knee
a. 'someone's knee'
b. 'a knee'
c. 'his/her (POLITE) knee' (cf. *bi-god* 'his/her knee')

Polite reference via inclusive

We do not find inclusivity appropriated for polite reference in the sample[11].

2.1.4.2 Politeness/Humbling directed at the first person

Politeness directed at the first person has two reflexes: the plural of majesty (akin to "royal *we*" in English) and the plural of humbling. Here, we will talk about both, even though the overall phenomenon is exceedingly rare, as Table 2.8 shows.

Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	6	0	0	4	0	0	1
Language	–	Persian	–	–	Guarijio	–	–	E. Kayah
		mā			*remé*			*ʔū*

Table 2.8: Recruitment asymmetries: Singular polite self-reference

Majesty/Humbling via plural

The plural of majesty is used for honorific self-reference in English (the "royal *we*"), but also in Thai, where the first person plural pronoun *jeu:ng* can only be used by royalty for

[11]Outside of politeness, the inclusive is also recruited for expressing solidarity with an addressee, or contempt towards an addressee. One example of the solidarity-inclusive is the English "nurses' *we*" or the "parents' *we*", as in (ia-b), but also attested in languages such as Lewo (Vanuatu, Polynesian), where the speaker is giving a speech to a large group but uses the inclusive dual to make it as if they are addressing a more intimate group as in (ii).

(i) a. (nurse to patient) How are **we** doing today?
b. (parent to child) Give **us** a hug.

(ii) Lewo (Early 1994: 114)
Tusi Wa ∅-pisa ke ∅-pa-n ***ita**-lua* *pae ∅-pisape.*
book holy 3sS-R.say TA 3sS-R.go-TR **1INCL**-two TA 3sS-R.say.that
'The Bible is telling us all today, saying that...'
Literally: 'The Bible is telling <u>us two</u> today, saying that...'

The positive politeness strategy above is to include the speaker for purposes of building solidarity. In Paamese (sg-du-pc-pl, with clusivity): INCL.DU forms for affectively addressing large crowd: "The speaker is effectively speaking to each addressee individually, and the use of the dual in such circumstances is more directly appealing to the individual in the large group" (Crowley 1982: 80). The same non-literal use of the dual inclusive pertains in at least Djambarrpuyngu (Corbett 2004: 224 citing Michael Cooke, p.c.) and in Tuvalu (Besnier 2000: 388).

As for contempt, Santali is reported to use the inclusive for such purposes, as (iii) shows. The same is reported for the Malayalam inclusive *nammaL* (India, Dravidian, Chandrasekhar 1970: 246).

(iii) Santali (Cysouw 2005: 221)
*...nãhãk'-**laŋ*** *ger-gitic'-gɔtme-a ar boge-te-**laŋ*** *thəyə-me-a*
...just-**1INCL.DU** bite-lie-2SG.O-IND and good-INSTR-**1INCL.DU** kick-2SG.O-IND
'... I shall just bite you that you lie there, and I shall give you a good kicking.'
Literally: 'I shall just bite <u>us</u> that <u>we</u> lie there, and I shall give <u>us</u> a good kicking.'

self-reference. In Kannada, the 1PL pronoun *naavu* is similarly used in lieu of 1SG *naanu* for asserting one's own majesty (Schiffman 1993: 38).

In contrast, plural can also be used for humbling oneself. We provide an example in Persian. (37) is an utterance made by a singular child to their schoolteacher, where the child is excusing themself from class early. The speaker explicitly avoids using 1SG *man*, instead using 1PL *mā*. Recall from (33) that Persian already allows honorific plural in the second and third persons; (37) shows that Persian allows this for all three grammatical persons.

(37) *khānom,* ***mā*** *emruz bāyad dah daghighe zudtar be-rav-**im**.*
miss 1**PL** today must ten minutes earlier SUBJ-walk-1**PL**
'Miss, we (HUMBLE) have to leave 10 minutes earlier today.' (Yousef 2018: 61, adapted)

Like Persian in this regard is Tafi (Ghana, Volta-Congo), where 1PL *blɔ* is used in lieu of *1sg mí* for self-humbling (Bobuafor 2013: 87).

Lastly, in Hindi, 1PL *ham* can be used instead of 1SG *mai* as either a plural of majesty or a plural of modesty (Bhatt 2012: 58-9).

Humbling via inclusive

Recall from (28) that the inclusive in Guarijio is used for addressee honorification, repeated below as (38a). However, the use of inclusive for politeness is symmetric, where both speaker and addressee use the inclusive. (In an asymmetric exchange, 1SG *ne* and 2SG *muú* are used instead.) In (38b), we see that the interlocutor also uses the inclusive *remé*, this time to refer to themselves instead of using 1SG *ne*. Here, the inclusive is directed towards grammatical first person, respectful because it indicates greater social deixis not by propping up one's interlocutor, but by putting oneself down.

(38) Guarijio symmetric honorific inclusive
a. *Muké-ru* ***remé*** *ku'írabi?*
carry-PAST.PASS **1INCL** short.while
'Did you (POLITE) carry it on your (POLITE) shoulder a while?'
Lit.: "Did <u>we</u> carry it on <u>our</u> shoulder a while?'
B. *Ka'í=**reme** muké-∅-**reme**.*
no=**1INCL** carry-PAST.ACT=**1INCL**
'No, I (HUMBLE) didn't carry it.'
Lit.: "No, <u>we</u> didn't carry it.' (Miller 1980: 197)

A handful of Malayo-Polynesian languages spoken in Indonesia also recruit the inclusive for self-humbling. These are Duri (Valkama 1995: 50), Galela (Shelden 1991: 166), and Sawu (Walker 1982: 11).

Humbling via fourth person

In Eastern Kayah Li (Myanmar/Thailand, Sino-Tibetan), the remote third person pronoun *ʔū* "may also be used as a humilific first-person pronoun" (Solnit 1997: 149, 184). In its non-social use, it is described as an indefinite pronoun, typically used for backgrounded referents, in contrast to the proximal third person pronouns *ʔa* and *lū*. However, all of *ʔū, ʔa* and *lū* are unmarked for number. It is unclear whether the proxy here is third or fourth person, but for the purposes of Table 2.8, we will list it as a fourth person. This is because fourth person pronouns are also number-neutral, and also typically serve as indefinite pronouns.

2.1.4.3 Politeness directed towards a group

At this stage, we have almost covered the entire space of logical possibilities, covering politeness directed towards first, second, and third persons, when these recipients are also singular individuals. The only remaining proxand is politeness directed towards groups.

Polite plural address via (inclusive and) higher number

In several of the languages covered above, there is a "ceiling" effect: the same form is used for honorific address, whether be it directed towards a singular or plural addressee; e.g. French *vous* is used for both instances.

Languages without this ceiling effect are only visible as such because they have more articulated number systems, allowing finer-grained number mismatches. Limbu (Bhutan/India/Nepal, Sino-Tibetan) is a language with morphological dual, and illustrates the recruitment of both plural and inclusivity. It makes number distinctions between singular, dual, and plural, along with a clusivity distinction, as in (39). van Driem (1987: 77) analyzes verbal agreement for 1INCL and 1EXCL via competition: *a-* expones the general first person, but is null in any verb form also containing the exclusive suffix *-ge*. This is because the category of exclusivity already entails first person, occurrence of an exclusive morpheme renders the first person prefix superfluous. (Harbour 2016: 107 arrives at the same conclusion for *-ge* as the exponent of exclusive.)

(39) Limbu pronouns and agreement (van Driem 1987: 25, 77, 80, 81)

	SG	DU	PL	agreement
1INCL	–	*anchi*	*ani*	*a*-V
1EXCL	*aŋga*	*anchige*	*anige*	∅-V-*ge*
2	*khɛnɛ?*	*khenchi*	*khɛni*	*kɛ*-V
3	*khɛŋ, khunɛ?*	*khunchi*	*khɛŋha?*	∅-V

Limbu features honorific inclusivity and honorific number: "In using the polite inclusive, the dual is used when addressing one person, and the plural when addressing two or more persons" (van Driem 1987: 221-2). The exclusive, marked with *-ge*, is not used as such.

(40) Limbu (adapted from van Driem 1987: 221)

a. *hn* ***a****-dzo:k-si-rØ* ***a****-ya:k-si?*
what **1INCL**-do-DU-PRES.GER **1INCL**-be-DU
'What are you (POLITE) doing?'
Literally: 'What are <u>we two</u> doing?'

b. *hn* ***a****-dzo:k-∅-lØ* ***a****-ya:k-∅?*
what **1INCL**-do-PL-PRES.GER **1INCL**-be-PL
'What are you (plural) (POLITE) doing?'
Literally: 'What are <u>we all</u> doing?'

Hence, when the proxand is polite address towards a group of individuals, the visible proxies involve higher number: plural when the group consists of two, but again plural when the group consists of more than two. (Polite reference towards a group of individuals, and polite-self reference towards a group of speakers, are unattested. Perhaps these proxands are possible, but there seems to be no visible reflex for them.)

Polite plural reference via higher number

In a language that possesses morphological dual, polite reference to two individuals may utilize the plural instead of the dual. This is Caddo (Oklahoma, Caddoan), where the plural

is used to refer to a group containing in-laws (41) (whether one or both of the group may be in-laws). Here, the plural suffix *-wa* is used, instead of dual *-wiht-*. (A corresponding example for polite *address* towards such a group is lacking.)

(41) *kúyt- 'a= **wa**= dih- hah? (=kúyt'awadihah)*
where- AGENT= **PL**= go- HAB
'Where are them two (POLITE) going?'
Lit.: 'Where are <u>they (more than two)</u> going?' (Chafe 1990: 66, adapted)

2.1.5 The negative evidence

Having covered all attested reflexes of politeness which span several ϕ-featural categories (person, number, and clusivity), we are ready to be explicit about what is unattested and cement a cross-linguistic asymmetry.

First of all, observe that the typology attests to the *lack* of mismatches in directions opposite to the ones we have already seen. For instance, no language recruits singular for plural honorification, as the hypothetical language hcnerF does: recruiting the singular pronoun *tu* for expressing politeness towards a group of addressees (42). Languages which have a grammatical option for honorifying groups use non-singulars instead, whether it is the plural or a number intermediate to plural, as we saw above in §2.1.4.3.

(42) hcnerF honorific singular (constructed)
As ***tu*** *le livre?*
have.PRES.2**SG** 2**SG** the book
Intended: 'Do you (plural) (POLITE) have the book?'

This asymmetry shows us that non-literal uses of *higher number* are attested politeness strategies, while non-literal uses of lower number are not. Towards one individual, a polite speaker marks them *as if* they were not singular: by adding others (resulting in the corresponding plural form) or by adding the addressee and others (resulting in an inclusive form).

As for person mismatches, there are no languages which recruit first person for addressee honorification as the hypothetical language kipuY in (43). Rather, there are several languages which recruit third person for addressee honorification, such as Yupik, repeated from (15).

(43) kipuY honorific first person (constructed)
*ner'-**gu**=am!*
eat-IND.**1**SG=again
Intended: 'You (POLITE) are eating again!'

(15) *ner'-**uq**=am!*
eat-IND.**3**SG=again
'You (POLITE) are eating again!'

Neither are there languages which recruit (exclusive) first/second person for referent honorification, as oddaC does in (44). Rather, there are several languages which recruit fourth person for referent honorification, such as Caddo, repeated from (98b).

(44) oddaC honorific first/second person (constructed)
*dikat- **ci/yah-** 'adih- hah?*
what- **1EXCL/2-** go- HAB

Intended: 'Where is he/she (POLITE) going?'

(98b) *dikat- **yi-** 'adih- hah?*
what- **IMPERS-** go- HAB
'Where is he/she (POLITE) going?'

This asymmetry shows us that non-literal uses of *higher person* are valid politeness strategies, while non-literal uses of lower person are not. Towards an addressee, a polite speaker marks them *as if* they were not second person by using a grammatically third person or fourth person form. Towards a referent, a polite speaker marks them *as if* they were not third person, by using a fourth person form. We saw this above in §2.1.4.1.

Hence, the relevant abstractions for politeness are those of *higher number* and *higher person*. In what follows, we will see that these abstractions persist in the face of extensive linguistic variation: whether a language offers multiple options for a polite pronoun (detailed in §2.1.6), or grammaticalizes multiple levels of politeness (detailed in §2.2).

2.1.6 Recruits in free variation

In a state of affairs where no one recruit is more polite than another, languages may enjoy multiple options for singular polite address. (Here, we go back to solely considering the proxand of singular polite address, as the other proxands are much rarer.) It is difficult to say whether these free variants embody steady states or changes-in-progress, what is important for our purposes here is that *all free variants swim only within the same circumscribed pool of ϕ-features established*, whether they offer optionality for different interactional settings or different stages of a language.

2PL, 3SG *for singular polite address*

Languages may have distinct honorification strategies operating in tandem, each reserved for different interactional contexts. Exemplifying this is Warlpiri (Laughren 2001), where honorific plural and honorific third person are both active. The strategy of honorific third person is used only if the interaction is taking place within a ceremonial context; for example, in initiation rituals, where boys are initiated into manhood. Within such ceremonies, the speaker addresses other participants with a 3SG pronoun. In contrast, the strategy of honorific plural is used for everyday contexts where politeness is required; for example, if the speaker is addressing anyone related by marriage, or anyone who was co-initiated with the speaker. Languages like Warlpiri are counted twice for this reason: once in the set of languages recruiting plural, once in the set of languages recruiting third person.

2PL, 3PL *for singular polite address*

In Slovenian (mainly spoken in Slovenia, Slavic), the 2PL pronoun *Vi* is used for honorific singular address, but Priestly (1993: 414-5) notes that 3PL *Onikanje* is also a valid substitution in some dialects[12].

Similar variation is reported for Bemba (Zambia, Bantu), where adults are addressed politely with either the 2PL or 3PL form (Heine & Song 2014: 132, citing Gregersen 1974: 53).

2PL, 1INCL *for singular polite address*

Duri (Indonesia, Malayo-Polynesian) offers two proxies for this purpose: 2PL *kamuq/komuq*

[12]Corbett (2004: 226 footnote 4) briefly notes that the 2DU pronoun *Vidva* can be used for honorific singular address in the written language, but no examples of this were given.

or 1INCL.PL *kitaq* (Valkama 1995: 50).

2PL, 4 *for singular polite address*

This variation can also occur between plural and fourth person. In Mongolian (Mongolia, Mongolic), 2SG *či* is used with intimates or well-acquainted social equals. On the other hand, 2PL *ta* is used for expressing social distance, as in (45b), whether it stems from age differences, or sarcasm, as is the case in this example.

(45) Mongolian polite plural (Brosig 2018: 106-7)

a. ***Či(n)**-ii ömön or-či=x=uu?*
2**SG**-GEN before enter-COMPLETIVE-FUT.PART-Q
'Can I defend [my dissertation] before you?'

b. ***Ta(n)**-iig xarin sons-loo.*
2**PL**-GEN but hear-PAST
'I did hear you (POLITE).'

When speakers do not wish to make any expressive stance regarding social distance, but still want to avoid expressing disrespect and maintain formality, the reflexive *öör-* is used without an overt subject pronoun. An example of such a context is interactions with strangers. This is illustrated in (46), where the reflexive makes a suggestion less insistent.

(46) Mongolian polite reflexive (Brosig 2018: 107-9)

a. ***Öör**=öö zav-tai=yuu?*
REFL=REFL.POSS time-COMPL=Q
'{Do you (POLITE) / Does s/he (POLITE)} have time?'

b. ***Öör**=öö bid(en)-tei xamt xözör togl-(x)=uu?*
REFL=REFL.POSS 1PL-COMPL with cards play-FUT.PART=Q
'Would you (POLITE) like to play cards with us?'

In Silacayoapan Mixtec (Mexico, Otomanguean), fourth person is recruited for both referent and addressee honorific: "the unspecified pronoun *a* is often used for respected referents, such as supernatural beings... it is also used for a respected second person" (Shields 1988: 408). In addition, plural is also used for honorification, as (47) shows for imperatives. (47a) with singular ∅ is perceived as a less polite imperative than (47b) with plural *ndó*.

(47) Silacayoapan Mixtec honorific plural

a. *nāhā* ∅*!*
come.IMP 2SG
'Come!'

b. *nāhā* ***ndó!***
come.IMP 2PL
'Come!' (Shields 1988: 340)

Lastly, as one may recall in §2.1.4.1 above, the Bear Lake Athabaskans described by Rushforth (1981:35-6) had the option of either using plural verbal agreement, or by using a fourth person subject to refer to in-laws.

2PL *for singular polite address; 4 for singular polite reference*

In Kambaata (Southern Ethiopia, Afroasiatic), for polite address, plural is recruited so that 2PL agreement is used (48a). For polite reference, it is not plural which is recruited so that 3PL

agreement is used, but fourth person agreement (48b).

(48) Kambaata (Treis 2008: 331-2, adapted)

a. *Á'nnu kánne ang-á-'nne barg-**iteenáni**-yan án ká'e*
2PL.NOM here hand-FEM.ACC-2PL.POSS add-**2PL**.ICO-DS 1SG.NOM there
kaa'll-áamm.
help-1SG.IMPF
'While you (POLITE) are helping here, I will help there.'

b. *Íssa qeh-óon ha'mm-ichch-ú dass-**éenno**.*
3PL.NOM chopping.tool-FEM.ICP enset.corm-SG-M.ACC chop-**IMPERS**.IMPF
'S/he (POLITE) will chop the enset corm with the qehúta.'

Diachronic perspectives on recruitment

Lastly, free variation amongst recruits can also be appreciated through a diachronic lens. It reveals assuringly recurring patterns, showing that diachronic options are circumscribed in exactly the same ways as synchronic options are. The recruits described below served the purpose of polite singular address.

Germanic. In Danish (Aalberse 2009: 69), 2PL *i* was once used, before moving on to the current recruit, the 3PL pronoun *De*. In German (Horst 2006), 2PL *ihr* was used throughout 1100-1950. Later, the 3SG pronouns *er/sie* appeared briefly as alternatives to *ihr*: "by the late 1500s, it became fashionable to use a new way of addressing someone when one wanted to be extra-polite and felt *ir* not to be enough" (p. 95). The current recruit, 3PL *Sie*, emerged around the 18th century. (It is unclear whether there is some linking principle between free variants and diachronic change. For instance, in German, once the recruitment of higher number 2PL and higher person 3SG were independently made available throughout the history of the language, this paved the way for 3PL to be recruited. I will merely acknowledge this possibility here because I do not have anything insightful to offer on the issue of what kind of factors were responsible for cementing particular choices.) Heine & Song (2014: 127, citing Simon 1997; Simon 2003) report that for a short period between the mid-18th to 19th centuries, the anaphor *dieselben*, literally 'the same ones', was recruited for plural polite address in place of 2PL *ihr*. Polite *dieselben* was not conventionalized to a full-fledged second person pronoun.

In Swedish (Dominika Skrzypek, p.c.; Clyne et al 2006), the 2PL pronoun *ni* and the 3SG pronouns *han/hon/hen* were distinct options for indicating some kind of social distance from an addressee: 2PL was derogatory while 3SG was respectful. A conscious reform leveled these social-deictic distinctions, leaving only 2SG *du* as the option for address. However, the younger generation has again co-opted 2PL *ni* for polite address, much to the ire of the older generation who still remember it being derogatory.

In Old Norse (Gumundsson 1973), recruitment of plural for politeness resulted in the loss of a grammatical dual. The plural pronouns in first and second persons, 1PL *vér* and 2PL *þér*, were re-appropriated for polite singular. New plural forms were innovated by re-appropriating the then-dual pronouns, 1DU *vit* and 2DU *þit*. This innovation remains in modern Icelandic as plurals, 1PL *við* and 2PL *þið*, and modern Icelandic does not have a morphological dual. A similar leveling occurred in the closely related language, Faroese (Faroe Islands, Lockwood 1975).

Early Modern English, spoken around 1600, had recruited the then-2PL pronoun *you* as a polite substitute for 2SG *thou*. The number distinction in the second person was leveled around the 1800s, so that in Standard English today, *you* is number-neutral.

Romance. In Spanish (Bentivoglio 2003: 177-8), the then-2PL pronoun *vos* was used for singular polite address, before being substituted by *vuestra merced*, which developed into the present-day polite pronoun *usted*. *Vos* reigned from the tenth century onwards (Lapesa 2000: 316-8). From the sixteenth century on, 3SG *él* also appeared as an option for singular polite address alongside 2PL *vos*.

Slavic. In Slovak, Slovenian, and Polish, the 3PL pronoun *oni* was the former option for singular polite address, before being ousted by the current option, 2PL *vy/wy* (Aalberse 2009: 69).

Dravidian. Brown & Levinson (1987: 293) observe that, throughout Old, Intermediate, through to Modern Tamil, the plural pronoun was recruited for honorification. Thus Old Tamil 2PL *nii-yir* (2SG *nii*) was recruited for singular polite address; in Intermediate Tamil the plural pronoun was repluralized with plural *-kaḷ* giving *niim-kaḷ*; in Modern Tamil *niin-kaḷ* is recruited for singular polite address (2SG *nii*). Parallel developments occurred in third person. Thus Old Tamil 3PL *avaar* (3SG *avaan*) was recruited for singular polite reference; in Intermediate Tamil the plural pronoun was repluralized with plural *-kaḷ* giving *avaar-kaḷ*; in Modern Tamil *avaar-kaḷ* is recruited for singular polite reference (3SG *avaan*).

Sinitic. In classical Chinese (Siewierska 2004: 219 citing Lee 1999), an individual of inferior status (typically women, servants or children) employs the 1PL pronoun for self-reference instead of 1SG when conversing with someone of superior status. The orthography also reveals a previous recruitment of fourth person. In a former stage of Mandarin, the characters for honorific address are either 您 *nin2* or 恁 *nen3*. The former is constructed from the combination of neutral second person pronoun 你 *ni3* and the heart/sincerity radical 心 *xin1*. Tellingly, the latter is constructed from the same heart/sincerity radical 心 *xin1*, but atop it sits the impersonal morpheme 任 *ren4*, 'any'. 恁 is now pronounced *nen4*, and currently has the uses of 'such', a distal demonstrative, 'how', and an associative marker.

2.2 Multiple-tier politeness systems

So far, we have only covered honorific systems featuring one level of politeness with a binary distinction between honorific and non-honorific. However, the empirical landscape is made much more complex with languages which carry out multiple recruitments throughout multiple levels of politeness, between non-honorific, honorific and high-honorific; I will refer to these as "tiered systems".

Strikingly, even in tiered systems, the same recruitment asymmetries emerge. Thus, tiered systems teach us that some recruits are more polite than others, and help us establish *clines* of politeness that individual recruits are situated on relative to each other. Thus, what seems like a source of delirious—or even deleterious—complexity at first glance is recasted into one of palatably informative complexity. Tiered systems provide a rich and nuanced source of typological variation, informing the final proposal in important ways.

§2.2.1 gives an introduction to tiered systems. §2.2.2 highlights tiered systems in Dravidian, while §2.2.3 highlights those of the Indo-Aryan. Indo-Aryan is particularly notorious for featuring several tiers which are morphologically specialized—seemingly—for honorification. Using diachronic data, I show that individual tiers are compatible with a recruitment-based explanation of their origins. Throughout, I show that the same recruitment asymmetries pertain, and that we can establish clines of politeness from them.

2.2.1 Tiered systems: An introduction

In this section I outline some instances of multiple recruitment. We will see that, in general, utilizing a person mismatch (e.g. recruiting 3SG for singular polite address) is a more polite move than utilizing a number mismatch (e.g. recruiting 2PL for singular polite address). In what follows, the notation $X \triangleleft Y$ indicates that X was recruited for a lower tier of politeness than Y.

2PL ◁ 3PL.

Consider Tagalog, which offers honorific plural and honorific third person plural as ranked options for singular honorific address: both 2PL *kayó* and 3PL *silá* are valid polite substitutions for 2SG *ikáw*. The former option is used for marking some social distance (e.g. from younger to older individuals, employee to employer), while the latter option is reserved for indicating "very great respect" (e.g. from subjects to monarchs, citizens to those in political office) (Schachter & Otanes 1972: 90-1).

Isomorphic to Tagalog are Ilocano (also known as Iloko, Northern Philippines, Malayo-Polynesian) (Rubino 1997: 52) and also Dii (Cameroon, Volta-Congo) (Bohnhoff 2010: 91). These 3-tier systems (non-honorific—honorific—high-honorific) are schematized below.

(49)

	2SG	◁	2PL	◁	3PL
Tagalog	*ikáw*		*kayó*		*silá*
Ilocano	*sika*		*dakayo*		*isuda*
Dii	*mí*		*ví*		*vó*
			polite address		

2PL ◁ 1INCL.PL.

Tukang Besi (Indonesia, Malayo-Polynesian) exhibits a tiered system of a slightly different type. 2PL *ikomiu* is used for singular addressee honorification. However, for a greater degree of respect toward a singular addressee, it is 1INCL.PL *ikita* which is used (50) (Donohue 1999: 114).

The same ranked system is described for Banggai (Indonesia, Malayo-Polynesian) (van den Bergh 1953: 43-4, translation my own). The 2PL form *ikomuju* is used for polite address, but the 1INCL.PL form *ikita* is reserved for addressing supernatural beings such as God and one's ancestors.

Such ranked systems are schematized below.

(50)

	2SG	◁	2PL	◁	1INCL.PL
Tukang Besi	*iko'o*		ikomiu		ikita
Banggai	*ikoo*		ikomuju		ikita
			polite address		

2PL ◁ 3SG ◁ 3PL.

In 19th century German (Manfred Krifka, Sarah Zobel, Nina Haslinger, p.c.), the choice of pronoun for singular polite address was differentiated according to social context, and resulted tiered system in (51). 2PL was used towards those in one's inner circle or in-laws; 3SG towards those outside one's inner circle, such as servants or distant acquaintances; and 3PL towards superiors.

(51) **19c German**

2SG	◁	2PL	◁	3SG	◁	3PL
du		*ihr*		*er/sie*		*Sie*

polite address

2.2.2 Dravidian

In most of Dravidian, we see the familiar pattern of recruiting plurals for honorifics, and reflexives (*albeit* plural reflexives) for high-honorifics. Below, we use Malayalam, Telugu, and Kannada to illustrate. (Within Dravidian, Tamil is an outlier. Tamil features only one option for politeness, which is formed by recruiting the plural.)

Malayalam. In Malayalam (Asher & Kumari 1997: 259), the 2SG pronoun *nii* is reserved for addressing intimates, while the 2PL pronoun *niŋŋaḷ* is used for polite singular address. This picture is familiar. However, the reflexive *taan-* is used socially as well, to address familiar friends. Its plural form, *taaŋkaḷ*, was used for polite singular address as *niŋŋaḷ* is used today. Today, *taaŋkaḷ* signals a highly formal conversation, a tier above *niŋŋaḷ*. For polite singular reference, there is also optionality between using either 3PL *avaar* or *taaŋ(kaL)*. Thus, we arrive at the following politeness cline for Malayalam as in (52).

(52) **Malayalam**

						polite reference		
				avaḷ		*avar*		*taaŋkaḷ*
2SG	◁	2PL	◁	3SG	◁	3PL ,		4+PL
nii		*niŋŋaḷ*						*taaŋkaḷ*

polite address

Telugu. Telugu (Krishnamurti & Gwynn 1985: 73, 76, 77) exhibits a tiered system similar to that of Malayalam: plural is recruited for the honorific tier for both second and third persons, while the plural reflexive is recruited for the high-honorific tier, but only in the second person. (The singular reflexive *tanu* is not used for honorification.)

(53) **Telugu**

						polite reference		
				avaLu/avanu		*avaru*		
2SG	◁	2PL	,	3SG	◁	3PL	◁	4+PL
niiwu		*miiru*						*taamu*

polite address

Kannada. The plural proxy of politeness is pervasive in Kannada: the plural is recruited for politeness for all numbers, persons, and genders (Schiffman 1983: 37-8). Thus, for singular polite address, 2PL *niivu* is used in lieu of 2SG *niinu*. For singular polite reference, 3PL *avaru* is used in lieu of 3SG.FEM *avaLu* or 3SG.MASC *avanu*. (Like English, plural makes no gender distinctions in the third person, and the third person plural can also be used in gender-agnostic contexts.)

For the additional high-honorific tier, it is the reflexive which is recruited (termed the "ultrapolite" tier by the author). The singular reflexive *taanu* 'oneself' is not honorific, while its plural counterpart *taavu* 'themselves' is used for both polite address and polite reference.

	Kannada									
								polite reference		
					avaLu/avanu		*avaru*		*taavu*
(54)	2SG	◁	2PL	,	3SG	◁	3PL	◁	4+PL
	niinu		*niivu*						*taavu*
		polite address							

Thus, Dravidian features 3-tier systems for both the second person and the third person (with the exception of Telugu having a 2-tier system for third person, as *taamu* is not used for polite reference). Plurals are more polite than singulars, and (plural) reflexives are in turn more polite than plurals.

2.2.3 The specter of Indo-Aryan

The pronominal systems of numerous Indo-Aryan languages feature distinct tiers of honorificity, specialized for both pronouns and verbal agreement. A typical language in this family makes the distinction between Non-Honorific (NH), Honorific (H), and High-Honorific (HH) in pronouns and agreement, creating a 3-tier system. Loosely, the NH tiers are used for intimates or for expressing contempt; H is a socially neutral tier used for acquaintances and equals; HH is used for those in socially respected positions like ministers, professors, or superiors. Hence, Indo-Aryan makes ternary or even quaternary honorificity distinctions. This is no matter, especially as we have already seen languages which make available multiple recruitment in §2.1.6 and §2.2.1. The challenge is that, honorificity distinctions are made in lieu of person/number distinctions, resulting in seemingly specialized honorific pronouns.

Here, I present diachronic evidence that each specialized tier does exhibit a history of recruitment, and from precisely the sources anticipated. The profusion of honorificity distinctions does not pose a challenge, as it is circumscribed to the pool of recruits we have already established.

2.2.3.1 Eastern Indo-Aryan

To start, I present tables illustrating the various tiers of honorificity in Eastern Indo-Aryan, as exponed by pronouns and verbal agreement. Throughout, I illustrate pronouns in nominative case. Depending on the language, there may be 2 tiers of honorificity (Non-Honorific *v.* Honorific; NH-H), 3 tiers (Non-Honorific *v.* Honorific *v.* High-Honorific; NH-H-HH), to 4 tiers (Non-Honorific *v.* Honorific *v.* Mid-Honorific *v.* High-Honorific; NH-H-MH-HH). Only singular pronouns will be presented and discussed in this dissertation. Plural pronouns are formed predictably atop the singulars with the addition of a suffix, and interested readers are referred to Wang 2025: §3.2 for their discussion.

While complexity of the data increases at this juncture, it is important to keep in mind that such complexity is reined in by the nature of recruitment. Complexity does not compound with the number of honorificity distinctions. We have already delimited the set of recruits to the plural, third/fourth persons, and inclusive. Tiered systems merely make multiple recruitments from this set, instead of making a single recruitment. Furthermore, agreement morphology is also restricted, given the limited ϕ-featural space available for recruits to occupy.

Bangla, Assamese and Magahi. Let us start with Bangla, illustrated in Table (55). It features 3 tiers in the second person (2NH-2H-2HH) and 2 tiers in the third person (3NH-3H). This is shown by (55), taken from David (2015).

(55)

Bangla	pronoun	agreement
2NH	*tui*	*-iś*
2H	*tumi*	*-o/-e*
2HH	*apni*	*-en*
3NH	*śe*	*-ē*
3H	*tini*	*-en*

Assamese and Magahi are isomorphic to Bangla, featuring 3 tiers in the second person, and 2 tiers in the third person. This is shown in Table (56), with Assamese being informed by Goswami & Tamuli (2003) and Magahi by Verma (2003).

(56)

Assamese			**Magahi**		
2NH	*tɒe*	*-A*	2NH	*tu*	*-en*
2H	*tumi*	*-a*	2H	*tu*	*-a*
2HH	*apuni*	*-e*	2HH	*apne*	*-thi(n)*
3NH	*xi, tai*	*-e*	3NH	*u*	*-ai*
3H	*tɛõ, tekhet*	*-e*	3H	*u*	*-thi(n)*

Throughout, we see that pronouns and agreement are differentiated only by honorificity, and not by person/number[13]. At first glance, it seems that no recruitment has taken place.

However, diachronically, grammaticalization paths traced from Sanskrit via Middle Indo-Aryan (MIA) show the current paradigms are the result of specific patterns of recruitment. If Indo-Aryan behaves like the languages considered so far, we predict that the only ϕ-featural sources for honorific pronouns are plural, third person, and fourth person for honorific address on one hand, and plural and fourth person for honorific reference on the other hand. The diachronic data shows that these predictions are borne out—the pool of recruits for the current-day honorific pronouns are indeed limited in precisely these ways. Here, Bangla data is given as a representative of the Eastern subfamily; Chatterji (1926: 935-7) documents the same changes for Assamese, Magahi, and Oriya.

In both 2nd and 3rd person, the Sanskrit plural was recruited for honorification. Sanskrit 2PL (pronoun/agreement **yūyam»tum hē/*-atha*) developed into the current 2H tier: *tumi/-o, -e*, while Sanskrit 3PL (**tē/*-anti*) developed into the current 3H tier: *tini/-en*. In contrast, the singular was not recruited for honorification. Sanskrit 2SG (**tvam»tūm/*-asi*) developed into the current 2NH tier: *tui/-ish*. The same pertains for third person, where Sanskrit 3SG (**sa(h)/*-ati*) is now 3NH: *śe/-ē*. Hence, what was a singular-plural opposition in Sanskrit has evolved into a NH-H opposition in both second and third persons, whereby plural was recruited for honorificity.

What of the HH tier? Such a tier was an innovation in New Indo-Aryan, and originated from the reflexive (Chatterji 1926: 847-9). Sanskrit **atman* 'self', later MIA **appane*, developed into Bangla 2HH *apni*. This reflexive falls within the umbrella of fourth person forms, being person- and number-neutral. These 2HH pronouns also take 3H verbal agreement *-en*, which, as the above has shown, originated from 3PL. This aptly combines two mismatches (plural, 3rd person) for the verbal agreement of the highest-politeness tier.

At this point, the recruitment data allows us to establish partial clines of politeness. As before, I use the notation $X \triangleleft Y$ to indicate that X is a recruit for a lower level of politeness than Y is. For systems like Bangla with 3 tiers in second person and 2 tiers in third person, the diachronic recruitment profile adheres to (57).

[13]This raises the good question of the *synchronic* interpretation of these systems; this question is addressed in Wang (2025), which argues that languages with and without specialized morphology for honorification can receive the same analysis.

Bangla

(57)

							polite reference		
					3NH *śe, -ē*		3H *tini, -en*		
					↑		↑		
	2SG	◁	2PL	◁	3SG	◁	3PL ,		4
	↓		↓				↓		↓
	2NH *tui, -iś*		2H *tumi, -o/-e*				2HH *-en*		2HH *apni*
			polite address						

Bangla is now cast in a familiar light: its politeness cline obeys the clines that we saw above in §2.2.1. The only differences are that the (57) involves a diachronic pathway instead of straightforward synchronic recruitment, and that (57) exhibits this for singular polite reference in addition to singular polite address.

Observe that the pool for addressee honorifics is larger than the pool for referent honorifics. Indeed, there are three tiers for address (2NH, 2H, 2HH) but only two tiers for reference (3NH, 3H). This is explained immediately: the recruitment pool for referent honorifics hits a "ceiling" sooner than the pool for addressee honorifics does. This is because 3rd person is a valid recruitee for honorific address, but 2nd person is not a valid recruitee for honorific reference. This is a robust asymmetry throughout Indo-Aryan and beyond: no language possesses more ϕ-featural means for referent honorification than for addressee honorification[14].

Bhojpuri and Maithili. Let us move on now to languages which feature an additional intermediate tier of honorificity in the second person. These are Bhojpuri and Maithili. I term the additional intermediate tier the Mid-Honorific tier (MH). The resulting system features 4 tiers in the second person (NH-H-MH-HH), and 2 tiers in the third person (NH-H). Table (58) illustrates their pronouns and agreement, where Bhojpuri data is taken from Lohar (2020) and Maithili from Yadav (1996).

(58)

Bhojpuri			**Maithili**		
2NH	*tē*	*-e*	2NH	*tō*	*-e/-ē*
2H	*tu*	*-a/-u*	2H	*tō*	*-əh*
2MH	*raua*	*-ĩ*	2MH	*əhã*	*-i*
2HH	*apne kā*	*-ĩ*	2HH	*əpne*	*-i*
3NH	*u*	*-ẽ/-i*	3NH	*u, o*	*-0/-əik*
3H	*uhã ka*	*-ĩ*	3H	*u*	*-əith*

What of the origin of the additional intermediate tier, 2MH *raua/-ĩ*? (Like the other Eastern languages, Bhojpuri and Maithili also descended from Sanskrit through MIA and the same diachronies for its current 2NH, 2H, 2HH, 3NH, and 3H forms apply.) The origins of this MH form are murkier, but Chatterji (1926: 847-8) speculates that *raua* originates from a 3rd person nominal, '(a/the) royal'. In that case, the source of the current 2MH form would be a 3SG form. Its agreement morpheme is shared with 2HH and 3H, which, in both cases, originates from 3PL agreement.

Thus, for systems like Bhojpuri with 4 tiers in second person and 2 tiers in third person, their politeness cline adheres to (59). The 2MH tier is boxed for easy identification, and its corresponding agreement morpheme *-ĩ* is circled.

[14]We are only making this claim for ϕ-featural means; as such, polite address/reference terms which are ostensibly nominals, like *sir* or *his lordship*, are discounted from this claim. Many languages including Indo-Aryan feature a wealth of such terms. As these terms do not involve recruitment of grammatical features, they are treated as part of polite registers, which are left for future work.

Bhojpuri

(59)

							polite reference		
					3NH *u*		3H *uhãka, -ĩ*		
					↑		↑		
2SG	◁	2PL	◁	3SG	◁	3PL,		4	
↓		↓		↓		↓		↓	
2NH *tē, -e*		2H *tu, -a/-u*		2MH *raua*		2MH, HH *-ĩ*		2HH *apne*	
		polite address							

The powers of recruitment are stretched to its limits with languages of the Bhojpuri type, which feature an intermediate-politeness tier in the 2nd person: 2MH *raua/-ĩ*. Recall that *raua* most probably originated from a 3rd person nominal, '(a/the) royal', a 3SG form. If we must fit 2MH into a politeness cline, then, it slots into the only remaining open position: the intermediate position between 2PL and 3PL.

(59) illustrates the sense in which a recruitment framework is stretched to *anticipated* limits, and no further. In a language like Bhojpuri without a clusivity distinction, there are four possible recruitees for singular polite address: 2PL, 3SG, 3PL, and 4. Bhojpuri saturates precisely this possibility space, and we do not find languages which exceed four tiers for addressee honorification. In (58), Bangla saturates three out of the maximum of four possibilities.

The recruits for verbal agreement reinforce the clines depicted. If honorific exponents are recruited, not innovated, we predict that tiered systems are left with limited choice for verbal agreement, and that certain agreement morphemes will be shared across tiers. This squeeze is accentuated also because there is no agreement morpheme available for a person-neutral reflexive, whence the 2HH pronouns originated. For singular polite address, the verbal agreement recruits available are 2PL, 3SG, AND 3PL; for singular polite reference, the only verbal agreement recruit available is 3PL. Indeed, we find the 3PL agreement morpheme shared across several tiers: in Bhojpuri, 3PL *-ĩ* is shared across 2MH, 2HH, and 3H; in Bangla, 3PL *-en* is shared across 2HH and 3H. The 2PL agreement morpheme is used exclusively with the lower-politeness tier originating from 2PL, the 2H tier. Hence, verbal agreement also establishes that 3PL is a more polite recruit than 2PL.

2.2.3.2 Central, Northern, Northwestern, and Western Indo-Aryan

Eastern Indo-Aryan exhibits the most complexity in tiered systems. Outside this subfamily, languages either make honorificity distinctions in only second person, or exhibit a lesser grammaticalization of honorificity (involving synchronic recruitment rather than diachronic pathways).

Central Indo-Aryan. Here, we consider Hindi (Shapiro 2003: 291) as a representative. It paints a picture very similar to Eastern Indo-Aryan. For singular polite address, we see a 3-tier system (2NH *tu*–2H *tum*–2HH *aap*). Again, the H tier developed from the plural and the HH tier from the reflexive. For singular polite reference, we see a 3-tier system (3NH *vah* — 3H *ve* — 3HH *aap*). *Aap*, with its reflexive origins, is found to approximate polite reference only in literary and archaic contexts, and is not in contemporary use as a referent honorific (Yash Sinha, p.c).

Hindi

(60)

							polite reference	
				3NH *vah*		3H *ve*	(3HH *aap*)	
				↑		↑	↑	
2SG	◁	2PL	◁	3SG	◁	3PL ,	4	
↓		↓					↓	
2NH *tu:*		2H *tum*					2HH *aap*	
		polite address						

Recall from §2.1.4.2 that 1PL *ham* can be used instead of 1SG *mai* as either a plural of majesty or a plural of modesty. Although this proxand of polite self-reference is not included in (60), but plural is recruited for honorificity in all three grammatical persons in Hindi.

Northern Indo-Aryan. This subfamily is also known as the Pahari ('hill') languages, are spoken throughout the Himalayan regions of the subcontinent. It consists of languages like Nepali, Garhwali, and Kumaoni. Since these languages do not feature politeness-related recruitment, but rather distinct registers, they are left for future work.

Northwestern Indo-Aryan. Panjabi, Hinko and Siraiki (Bashir & Conners 2019: 189) feature a simple French-like system, with honorific plural in the second person. Instead of using 2SG *tü* (common to all three languages), 2PL is used for indicating social deixis, along with corresponding 2PL verbal agreement. This is *tusī*, illustrated for Panjabi in (61) (*tussī* in Hindko; *tussā* in Siraiki).

(61) ***tusī*** *kitth-õ* *d-**e*** ***o*** *ḍākṭar sāb?*
2PL where-ABL GEN-**PL**.MASC be.PRES.**2PL** doctor sir
'Where are you (POLITE) from, doctor?'

For Panjabi, the authors give an isolated example for plural being used for singular polite reference as well (Bashir & Conners 2019: 228); however, it is unclear whether this extends to Hinko and/or Siraiki.

Western Indo-Aryan. Western Indo-Aryan languages have innovated a clusivity distinction (absent in Sanskrit and the most of non-Western Indo-Aryan). Strikingly, across most of this subfamily, the reflexive **atman* from Sanskrit developed into the inclusive plural, and in some cases, also into the 2HH pronoun. This diachronic affinity between the developmental pathways of the inclusive and the high-honorific categories is discussed further in Chapter 3. For now, we will only discuss Marathi, which recruits the form *apəṅ* for politeness, a form which is ambiguous between inclusive plural and reflexive (62).

(62) Marathi pronouns and agreement (Dhongde & Wali 2009: 48-50)

Marathi	SG	PL	agreement
1EXCL	*mi*	*amhi*	-
1INCL	—	**apəṅ**	-
2	*tu*	*tumhi*	-
reflexive empathic	*swətəha*	**apəṅ**	–

Honorificity is not grammaticalized, so that there are no NH/H/HH tiers in Marathi. For singular polite address, both 2PL *tumhī* and *āpaṅ* are used in place of 2SG *tu*. For plural polite address, *āpaṅ* is used in place of 2PL *tumhī*. Traditionally, the inclusive was used for a higher level of politeness than the plural was. Although this use of the inclusive is dying out (Pandharipande 2003: 771), we see that it is possible for the plural to be recruited for a tier of politeness (2H) lower than that of the inclusive/fourth person (2HH). This forms the Marathi politeness cline as in (63), familiar from Tukang Besi in (50), except that high-honorificity is expressed via a form ambiguous between inclusive and fourth person.

(63) **Marathi**

2SG	◁	2PL	◁	4/INCL
tu		*tumhī*		*āpaṅ*

polite address

2.3 Moving on: Explicanda and Desiderata

2.3.1 Explicanda: asymmetries and sub-asymmetries

The typology has shown us that the pool of recruits for politeness is consistently limited to the following.

From §2.1, the languages with single-tier, binary systems (non-honorific *vs.* honorific) showed us that the proxies for all three proxands considered thus far—polite address, polite reference, polite self-reference—are limited. They are combined in Table 2.9.

Singular polite address								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	125	27	12	7	29
Language	–	–	–	French	Fehan Tetun	Italian	German	Ainu
				vous	*ita*	*Lei*	*Sie*	*an*
Singular polite reference								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	0	0	0	28	8
Language	–	–	–	–	–	–	Wolaytta	W. Apache
							ʔetaa	*ko-*
Singular polite self-reference								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	6	0	0	4	0	0	1
Language	–	Persian	–	–	Guarijio	–	–	E. Kayah
		mā			*remé*			*ʔū*

Table 2.9: Total recruitment asymmetries

From §2.2, the languages with tiered systems (making more distinctions than binary non-honorific *vs.* honorific) showed us that the pool of recruits for honorific address and reference are identically limited, *and* that some recruits are more polite than others. Here, we will establish some partial clines before assembling them into a total cline. We are unable to establish a cline for polite reference, however, since there is no language in the dataset which makes more than a binary distinction in the third person.

From Indo-Aryan systems, we have seen that fourth person was recruited for higher tiers of politeness (2HH), whereas plural was recruited for lower tiers of politeness (2H). In languages like Bhojpuri and Maithili, third person was recruited for an intermediate tier (2MH). Keeping in line with typological trends, singular is never recruited for politeness, and hence developed into a non-polite tier (2NH). This is schematized in Table 2.10. The subcline 2PL $\triangleleft$ 3 is also supported by 18th century German, which showed us that the 3PL recruit, being used for superiors, is more polite than 2PL, used towards intimates.

Lastly, languages which make clusivity distinctions, like Tukang Besi and Marathi, showed us that the inclusive is used in contexts warranting more politeness than plural, as Table 2.11 schematizes.

2.3.2 Desiderata

We want the recruitment asymmetries and its corresponding clines to have theoretical reflexes; to line up with another cline independently provided by the grammar. The account must not only

Recruit for...	2PL	◁	3SG	◁	3PL,	4
	2H pronoun, agreement		2MH pronoun, agreement		2HH agreement	2HH pronoun

Table 2.10: Singular polite address cline from Eastern Indo-Aryan

Recruit for...	2PL	◁	1INCL
	Marathi *tumhī*		Marathi *āpaṅ*
	Tukang Besi *ikomiu*		Tukang Besi *ikita*

Table 2.11: Singular polite address cline with inclusive

explain why certain recruits are attested and others unattested, but also why certain recruits are more polite than others.

The recruitment asymmetries established beg a specific explanation: what kind of natural class do the various attested proxies of politeness belong to? They are otherwise a heterogenous bunch, crosscutting the ϕ-feature categories of person, number, and clusivity. Furthermore, what property underlies their unattested counterparts so that they are never recruited? Answering these questions will allow us to understand the typological asymmetries for both single- and multiple-tier systems established.

The hypothesis that we will explore in the next chapter is that proxies of politeness belong to the natural class of *semantically unmarked elements*: I define these as semantic defaults carrying little to no specification in meaning, allowing polite speakers to be less direct. To do so, we must establish and diagnose the property of being a *semantic default*, and embed it within a theory of ϕ-features. It is to this task that we turn in the next chapter.

Chapter 3

On the way, a kindred spirit Semantic Defaults

The previous chapter established an empirical landscape with the following landmarks:

- Recruitment displays robust typological asymmetries. (§2.1.5)

  Plural, third person, and fourth person, and inclusive are phi-featural proxies for politeness, to the exclusion of singular, first person, and second person.

  This asymmetry pertains for single-tiered systems and for multiple-tiered systems. Warlpiri examples of plural and third person proxies for singular polite address are repeated below.

  (1) Warlpiri honorific plural

  *ngana-ngku-**nyarra** ngarrurnu yarlpurru-pardu?*
  who-ERG-**2PL**.OBJ told coinitiate-DEF
  'Who told you (POLITE), (my) agemate-brother?'
  Lit.: 'Who told <u>you (plural)</u>?' (Laughren 2001: 210-1, adapted)

  (2) Warlpiri honorific third person

  *Nyampu-rla-juku mayi ka-**Ø** karrka?*
  here-LOC-still Q PRES-**3**.SUBJ locate
  'Are you (POLITE) still living here?'
  Lit.: 'Is <u>s/he</u> still living here?' (Laughren 2001: 214-5)

  Yet, Warlpiri—or any other language—does not use a first person proxies for singular polite address. While this is a logically possible option, it is not attested. Hence, languages like iriplraW in (3) do not exist.

  (3) iriplraW honorific first person (constructed)

  *Nyampu-rla-juku mayi ka-**rna** karrka?*
  here-LOC-still Q PRES-**1**.SUBJ locate
  'Are you (POLITE) still living here?'
  Lit.: 'Am <u>I</u> still living here?'

- Recruitment displays a politeness cline—some recruits are more polite than others. (§2.2)

  For singular polite address, second person plural is the least polite option, while third person plural, fourth person and inclusive are the most polite options.

This leads us to the following questions:

- What is the social-anthropological reflex of these pervasive recruitment asymmetries?

  Claim: The reflex is identified as *negative politeness*, building on Goffman's (1967) notion of *face-work* and Brown & Levinson (1978). A polite speaker deploys *negative politeness strategies*, which are avoidance-based behaviors. The overall effect is that of avoidance or distancing, an effect that lies at "the heart of respect behavior" (ibid: 129).

- What, in turn, are the grammatical reflexes of this social-anthropological pressure? In other words, what grammatical feature makes certain ϕ-featural values suitable proxies?

  Claim: Proxies are all *meaning-poor*, in the sense that they are less informative for some dimension of meaning, being number-neutral, person-neutral, or both.

  Claim: To use a meaning-poor form *is* to avoid responsibility for the information conveyed. This is the linguistic reflex of negative politeness, well-suited as an avoidance strategy.

Roadmap

This chapter will describe and explain how the ϕ-featural expressions of politeness from Chapter 2 are recruited as such. We first explicate the nature of politeness itself, linguistic or non-linguistic, with help from the anthropological literature, establishing the relationship between negative face, negative politeness, and avoidance in §3.1. To understand the linguistic reflexes of negative face, I take us into the realm of a completely different linguistic phenomenon: semantic defaults. These are ϕ-featural values which are person-neutral and/or number-neutral. I propose various diagnostics for semantic defaulthood in §3.2 (being careful throughout to distinguish it from morphological defaulthood). We establish a surprising convergence: the ϕ-features recruited for politeness are precisely the ones which pass tests for being semantic defaults, based on English data. §3.3 establishes the same set of semantic defaults but with typological scope.

3.1 Social reflexes of politeness: Face and indirectness

3.1.1 Face and face-work: Goffman

Goffman (1967) was concerned with the "sociology of occasions": "the countless patterns and natural sequences of behavior occurring whenever persons come into one another's immediate presence". He identified *face* as the relevant primitive[1]: a public image of the self that the individual seeks to project onto others in social interactions, delineated in terms of approved social attributes. Since individuals are multifaceted beings, they can have many faces depending on the social situation they find themselves in: one projects a different face when interacting with superiors, colleagues, casual acquaintances, strangers, family, or intimates. Face is the ultimate example of a social construct: it exists only in the presence of others; it is irrelevant in contexts of isolation. Face is why we clean the house for guests; why people make a show of fighting for the check in restaurants; why we refrain from checking our mobiles when in a conversation. In the absence of these behaviors, we lose face: our public image (as tidy host, generous friend, or attentive listener) is damaged otherwise.

Face-work is what co-operative individuals do in interactions: maintaining the face of themselves and others. One can be defensive, saving his own face, or protective, saving others' face. Hence, natural sequences of behavior in social settings are boiled down into a set of face-related considerations, and sequences of behavior, wherever they occur, can be given a syntax in these terms.

The face that individuals project depends on the social situation they are in; this can be seen when they carry out face-work in the presence of others. Defensive face-work is at play whenever one maintains their own face, actively avoiding situations where it might be threatened. Let's suppose that Nigel has forgotten to clean his house for his New Year's guests. He can carry out defensive face-work for himself by apologizing and exaggerating the mess, or hiding the mess in a private room closed off to guests: *I'm really sorry, it's really a big mess in there.* Protective face-work is at play when one maintains others' face for them. Thus, his guests might indicate that they do not mind the mess, and reply along the lines of *That's alright, we have pets at home!* This is all to preserve Nigel's face since an explicit acknowledgement of the mess entails a social threat: exposing Nigel as an imperfect host who does not give his guests their ceremonial due.

Respectful and polite behaviors constitute a major class of face-work. A polite speaker ensures that others are treated in the way they are due; not doing so is a potential double offensive: it harms the speaker's own face because it indicates that the speaker is some anti-

[1] While Goffman is credited with integrating the notion of face into theories of human interaction, the notion originated from Chinese culture, and is still highly salient in Chinese society today.

social renegade, and it harms his interlocutor's face because it indicates that they are not worthy of some socially accepted treatment. These non-polite situations do not even have to involve explicit social institutions like differences in rank. For instance, it is considered very impolite to barge into someone's office when the door is closed, whether that someone is a colleague or a superior. This is because the act of proffering space is normally taken to be a symbolic means of appreciation for another individual's autonomy. To violate this is to indicate that you are boundlessly entitled to their time and energy, that you deem their time and energy to be worth less than your own.

When any kind of face-work is successful, the self-image projected by each individual is allowed to prevail. The social role they have chosen for themselves is temporarily acknowledged and accepted by other individuals, for the purposes of the social setting they are situated in. Such mutual acceptance merely constitutes a social lubricant, not necessarily a heartfelt solidarity and empathy. It is the willingness "to give temporary lip service to judgments with which the participants do not really agree." For instance, Nigel's guests might be very tidy and be inwardly disgusted at the mess, but, outwardly, they comply with the social ceremony of excusing the mess so as to let everyone enjoy the New Year's gathering without histrionics.

As hinted, this "temporary lip service" often characterizes "polite" situations which necessitate the use of certain mismatched linguistic forms. So-called polite forms are used not because of a genuine attitude of deep reverence, but because of a willingness to comply with whatever social conventions are already in place. Thus, presence of expected "polite" forms merely mark the fact that a social relationship has not changed since the last interaction. Nigel might not necessarily respect a social superior (a boss, elderly relation, etc.) but he will continue to use the polite pronoun with them; it is merely a means of getting through an interaction, its social lubricant. Certainly, social relationships can change depending on a variety of subjective personal factors. If Nigel switches to the non-polite pronoun with an individual, he intends to mark the fact that their relationship is changing; for instance, that they have become intimates since their last interaction[2].

3.1.2 Onto language: The facework of politeness

Brown & Levinson (1987) focused on face-work as it pertained to polite behaviors, dividing the Goffmanian concept of face into positive face and negative face. Two types of polite behaviors follow: positive politeness and negative politeness. They propose that each has distinct linguistic reflexes, basing their investigation of politeness on three unrelated languages and cultures: English, Tamil (India, Dravidian), and Tzeltal (Mexico, Mayan).

Like Goffman, Brown & Levinson recognize that to interact in any way is to introduce *face-threatening acts*. Their framework lays out the various redressive efforts that mitigate the impact of face-threatening acts, and the choice of redressive effort varies along their proposed factors of power difference, social distance, and rank of imposition. Since these are sociological factors outside the scope of this dissertation, I will not cover them in detail here. Rather, the focus is on the grammatical manifestations of redressive efforts, regardless of the social circumstances that triggered them.

Negative face and negative politeness

Negative face is the desire to be unimpeded: the recognition that every individual has a separate autonomy that is sacred and should not be impinged upon lightly. As Brown & Levinson (1987: 61) define it, it is an individual's "basic claim to territories, personal preserves, rights

[2]Brown and Gilman (1960: 111) observe that, in French, "mountaineers above a certain critical altitude shift [from mutual *vous*] to the mutual T [*tu*]... this is the point where their lives hang by a single thread."

to non-distraction – i.e. to freedom of action and freedom from imposition". Often, negative politeness strategies are characterized by avoidance behaviors. In Goffmanian terms, respecting another's negative face is to give them their ceremonial due: to not barge in on a closed office door, to not be entitled to their time and energy. This maximizes autonomy to the addressee and minimizes any potential obstruction that the speaker imposes.

The ultimate avoidance strategy is to avoid co-presence entirely. We have already seen this in Chapter 2 §2.1.4.1 with in-laws, where only honorific reference is found, to the exclusion of honorific address. In fact, "avoidance relations" in several societies are explicitly documented by anthropologists and ethnologists. In several communities of Aboriginal Australia, direct eye contact and/or physical contact with certain kin (typically in-laws) is strictly forbidden. If there must be interactions between avoidance relations, an intermediary is typically required. This can also be seen in the languages where the avoidance register is characterized by straight avoidance of a pronoun/proper name. Since these do not have grammatical reflexes *per se*, they were not included in the empirical findings in Chapter 2, but such findings must be considered by the present proposal.

Of course, avoiding co-presence and recruiting an intermediary are often impractical or impossible strategies. When contact must be made, there are strategies which mitigate potential threats to negative face. These are linguistic strategies which attempt to minimize the speaker's imposition onto their interlocutor: apologizing for imposing on the other, circumlocutions, hedging, and being vague. In Goffmanian terms, these constitute protective face-work.

Certain speech acts are fertile grounds for deploying these strategies. In particular, requests and imperatives are transparent threats to an interlocutor's negative face: wanting something costs nothing, but giving or granting it costs much more. Thus, they are typically not phrased as bare imperatives (*Pick me up from the airport!*), but rather as apologies (*I'm sorry to bother you, but...?*), circumlocutions (*Well, if you're around, I wonder you might have the time to...*), hedging (*Perhaps you could...*), and being vague (*Do you have errands to run near the airport by any chance?*). Even though these careful rephrases, as hints, constitute Gricean Quality violations, they enable the interlocutor to refuse. Thus, the interlocutor is given an out, as deploying negative politeness strategies adds a layer of deniability.

Deniability is crucial to negative politeness: it ensures that the interlocutor's negative face is preserved even if their autonomy has been briefly punctured by a request, or by any other engagement. A deniable proposition is considered less face-threatening, as it makes it possible for the interlocutor to ignore the speaker's imposition. For the interlocutor to say *no* to the thinly veiled request *Do you have errands to run near the airport by any chance?* is much easier than to say *no* to the command *Pick me up from the airport!*. By using the former, the speaker thus gives the interlocutor the option to easily decline, thereby freeing themselves from any responsibility to the speaker. The interlocutor's negative face is maintained in a pristine, pre-request state, as if they were never imposed upon. A passage from Goffman (1967: 30) about the hinted language that is contained in such polite circumlocutions is particularly helpful here:

> The rule regarding this unofficial kind of communication is that the sender ought not to act as if he had officially conveyed the message he has hinted at, while the recipients have the right and the obligation to act as if they have not officially received the message contained in the hint. Hinted communication, then, is deniable communication; it need not be faced up to.

Deniability is thus a desideratum for the final analysis of polite pronouns.

Positive face and positive politeness

Positive face is the desire to be approved of: that one's own face is appreciated by, and desirable to, other individuals. Positive politeness strategies, then, are those acts which reinforce an individual's positive face: giving compliments (instead of belittling or insulting), expressing agreement or validation (instead of disagreement or challenges), being interested (instead of interrupting or changing the subject).

Positive politeness does have linguistic reflexes; for example, they are used in the innovation of some items in polite registers; there will be some discussion of "lexicalized honorifics" in §5.3.3. However, a full discussion is outside the scope of the current work and is left for future research. Our focus now is on the grammaticalized reflexes of negative politeness.

3.1.3 "Indirectness"

Informally, we formulate a pragmatic rule, the Taboo of Directness in (4), to capture the effect of politeness-as-avoidance.

(4) Taboo of Directness (ToD):
Don't use forms which are direct!

To violate this maxim is to suffer a small social death, in accordance with what happens when negative face is threatened. The formulation of "directness" is deliberately left definitionally fuzzy at this stage, as it only reflects that what individuals in a polite social situation are sensitive to is *negative politeness*.

The expression of politeness, then, is to express pretense away from a situation that directly refers to an individual, to avoid them via linguistic means. The individual is singular, but they are marked *as if* they were plural; the individual is a second person, but marked *as if* they were third person; the individual is a discourse participant, but marked *as if* they were not (as fourth person); the individual is an addressee, but marked *as if* they were part of an author-addressee dyad (as inclusive). These strategies are all ways of diminishing the direct imposition deployed on an individual.

Pretense is useful for politeness because it grants the interlocutors deniability. One may also ask to what extent is this pretense conventionalized, so that interlocutors in a formal setting do not really buy into the pretense, but merely go along with it to avoid flouting social convention. Whichever it is, note that the conventionalization option is also a polite behavior. Not going along with it is not polite, as exempting oneself from a social convention is an arrogant thing to do.

What about positive politeness, which operationalizes approval instead of avoidance? We will see that the empirical findings from Chapter 2 are better understood as "avoidance mechanisms", but I do not preclude the possibility that there are items in polite registers which may derive from "approval mechanisms". These are left for future work.

The analytical challenge is to explicate the linguistic means responsible for these avoidance strategies, the mechanism that is the linguistic counterpart to (4). We will identify what the recruiting mechanism, situated in the core grammar, is sensitive to. §3.2 will relate the notion of face-work and negative politeness back to the empirical findings from Chapter 2.

3.2 Diagnosing semantic defaults

Our object now is to relate the Taboo of Directness in (4) to something in the core grammar. What is an example of a linguistic feature that would satisfy the Taboo of Directness? One

non-phi-featural example comes to mind: the passive voice, which allows speakers to leave the agent of an event unmentioned, as opposed to using the active voice (whether the speaker is intentionally trying to be vague, or simply does not know; while polite passives exist, we will not give a full-blown discussion of them in this dissertation).

Phi-features can also relate to the the Taboo of Directness. We eventually cash out "indirectness" via the notion of *semantic markedness* (e.g. Sauerland 2008): a semantically unmarked element is compatible with a wider range of contexts, by dint of being unspecified or underspecified in its meaning. As a result, when a context does not supply, specify, or require a particular value, the semantically unmarked element steps in, being a *default* or *neutral* value. There may be several symptoms of semantic unmarkedness, but the symptom relevant for us is semantic defaulthood. Hence, we will turn away from politeness phenomena for a while, focusing on the completely different grammatical phenomenon of semantic defaulthood.

We define *semantic defaults* are those featural values which are neutral with respect to other featural values in the same category. Hence, the default for the category of number (spoiler: plural) is *number-neutral*, while the default for the category of person (spoiler: third, fourth person) is *person-neutral*. However, most contexts are not amenable to defaulthood or neutrality. To diagnose defaulthood, then, I identify several "special" contexts which do allow the emergence of default or neutral forms, where the context is unable to demand, or supply, a specific feature value.

- *Contexts of epistemic uncertainty*:
  When the speaker of an utterance is not certain about the feature value that holds in a context, which feature value is realized?
- *Contexts of arbitrarity*:
  When an utterance features a non-referential subject (e.g. directed towards no particular individual because it is intended to be prescriptive, or for the general public), which feature value is realized?
- *Mixed quantification*:
  When a universal quantifier, conjunction, or disjunction ranges over multiple feature values in the same category, which feature value is realized as the bound variable?

In contexts of epistemic uncertainty, a specific feature value is simply unknown and choosing any one destroys epistemic uncertainty. In contexts of arbitrarity, non-referentiality is the goal and a specific person value is its very antithesis; all existing feature values are too specific. In scenarios with mixed quantification, no existing feature value is specific enough.

Other than these special contexts, imposter phenomena will also be important. Imposters are third-person expressions which can take on first or second person reference (e.g. Collins & Postal 2012, Podobryaev 2017), such as the third-person DP *my one and only* which may refer to the addressee.

- *Shapes of imposters*:
  Which feature value(s) can pose as another, for stylistic or rhetorical purposes?

We will eventually see that the meaning profile for feature values are *asymmetric*, so that certain feature values are more specified than other feature values. For person, first/second persons (participant persons) are more specified than third/fourth persons (non-participant persons). For number, singular is specified than plural, in both nominal and pronominal domains. We do not find evidence for symmetric meaning profiles, where all feature values are equally specified. This asymmetry will justify the privative system that is proposed in §4.1.2.

The epistemic uncertainty and mixed quantification diagnostics are modified from Sauerland's (2008) discussion on semantic markedness (although he terms mixed quantification "dominance"). In our discussion of contexts of arbitrarity, we draw data from Malamud's (2006) dissertation, which shows a range of pro-forms being used for this purpose. Both works are cross-linguistic in their aim as this dissertation is; I mainly provide additional evidence for their findings regarding person and number, but provide novel evidence regarding clusivity.

I observe that these special contexts are all contexts where *Maximize Presupposition!* (5) vacuously applies. There is no "maximal" presupposition suitable, as there is not sufficient information provided by the context. This is identified as the formal property unifying these special contexts.

(5) *Maximize Presupposition!* (MP!, Heim 1991)
Choose the strongest presupposition compatible with what is assumed in the conversation.

For instance, when cardinality is unknown, it mandates the speaker to use the number-neutral form because that is already the form with the strongest presupposition compatible. Using a number-specific form in this case would result in presupposition failure.

Importantly, the setup of these diagnostics is completely independent of politeness. This shows us that the recruiting mechanism is not sensitive to politeness itself; rather, it is sensitive to semantic underspecification. X can only be a proxy for Y if X is semantically unmarked relative to Y. Within a phi-featural category, *it is the proxy's feature value which has a wider meaning relative to its proxand's.* Even when politeness is not a factor, the recruited feature value is suitable for use in a wider range of contexts. Particularly, when discussing imposters, it is important to keep in mind that imposters are licensed by stylistic contexts that do not contribute a main effect of "politeness", that of increasing social distance between speaker and interlocutor.

Most of these diagnostics will be illustrated via English, as the evaluations of defaults and impostering are most convincing with negative evidence. However, for featural values which English does not illuminate because it lacks the relevant morphological distinctions (*e.g.* clusivity), typological data steps in. Since most typological data has been culled from descriptive grammars, however, negative evidence will be sparser than positive evidence, which is unfortunate. Given practical constraints, I cannot present language-specific results for these diagnostics in all 150 languages of the typology. Hence, the universality of the following results is a stipulation. However, it is also important to note that the typological method provides continuous opportunities for falsification, and it is notable that such falsification has not happened yet—not for the shapes of polite pronouns or semantic defaults.

Little of the final analysis is novel: there is no grammatical mechanism or feature that is intrinsically about politeness. The main innovation lies in identifying an established socio-anthropological notion (negative politeness) with an established formal semantics notion (semantic markedness) as crucial, and connecting them together. We show that the linguistic expression of politeness is expressly parasitic on semantic defaults, and will give a principled explanation of this relationship.

3.2.1 Manners of markednesses

Before we bridge indirectness with semantic markedness, let us consider an alternative bridge between indirectness and *morphological* markedness. What are the differences between the two notions of markedness, and why should semantic markedness the relevant notion for our purposes? Below, we consider morphological markedness, and burn this potential bridge.

A semantically unmarked element is said to have default or neutral interpretations (e.g. Sauerland 2008). Given a pair of values in the same category (e.g. SG versus PL in the category of number), the less marked element is compatible with a wider range of contexts because it carries a weaker meaning than its marked counterpart. In other words, a semantically unmarked element will more permissive in the range of possible interpretations compared to their semantically marked counterparts.

The concept of semantic markedness is independent from that of morphological or syntactic markedness, as they have complementary tasks. Morphosyntactic markedness governs *forms*, while semantic markedness governs *meanings*. However, across these distinct notions of "markedness", some form of extra complexity is involved for the marked member(s) (Haspelmath 2006: 26), although "complexity" will have different reflexes[3][4].

Reflexes of morphological complexity are several, as the study of morphological markedness constitutes a truly enormous body of literature, since its investigation has been ongoing since the inception of the field (tracing back to Trubetzkoy and Jakobson, then Greenberg 1966; Croft 2003; Harley 2008; *a.o.*). Some symptoms of morphological markedness are that unmarked members tend to: i) make fewer paradigmatic distinctions; ii) be less distributionally restricted (due to being targets of Impoverishment, or due to being an elsewhere form with no specification), and iii) have null exponence. We will briefly illustrate these diagnostics with English number.

We can illustrate number of paradigmatic distinctions with gender marking on third person pronouns. It is common that gender is only marked in the singular, and not in the plural (e.g. English *he/she/it vs. they*). Sometimes, this is explained in terms of Impoverishment rules targeting the plural, such as "gender is deleted in the plural" (Harley 2008). With regards to null exponence, singular is said to be morphologically unmarked, as it is often null compared to the plural (e.g. English plural/singular *-s/0*).

Meanwhile, *semantic* markedness is commonly assumed to only have one symptom (Jakobson 1971; Croft 2003; Silverstein 1976; Sauerland 2008): the unmarked pair member is able to refer to more types of entities than its marked counterpart. We can illustrate this with number marking in English, where plural is said to be semantically unmarked. In a context of epistemic uncertainty, where the speaker does not know how many hedge fund managers are in that day—then the speaker is unable to use the singular as in (6a). Rather, the speaker must use the plural, as in (6b).

(6) a. #I just started here and I haven't familiarized myself with everyone's schedules, but don't let yourself bump into **the hedge fund manager** on the way out.
b. I just started here and I haven't familiarized myself with everyone's schedules, but don't let yourself bump into **the hedge fund managers** on the way out.

The felicity of (6b) indicates that the plural is semantically unmarked, as it can range over both singularities and pluralities. A common idea is that singulars contain only atoms in their extension, while plurals contain both atoms and pluralities, as the toy semantics in (7) indicates. While the singular carries with it a restriction that the cardinality of its referent is one, the plural carries no restriction at all, being number-neutral. Our analysis of number will be very similar. Although (7) implements these restrictions via presupposition, we will not commit to

[3]Yet another markedness is syntactic markedness, where syntactically marked categories trigger exceptional syntactic behavior; for instance, plural triggers omnivorous agreement (e.g. in Georgian, Nevins 2011), is susceptible to ϕ-neutralization (e.g. Long Distance Agreement in Basque, Etxepare 2006) and is also more susceptible to agreement errors (e.g. Eberhard 1997; Tucker, Idrissi & Almeida 2015).

[4]There are certain environments where it is unclear whether morphological or semantic markedness is operative: disjunction (Pullum & Zwicky 1986) and infinitives. Since disjunctive phrases or infinitives are not attested proxies of politeness, I shall put these environments aside for now.

presupposition.

(7) A semantics for number (Sauerland 2008, *a.o.*)

a. $[\![\text{SG}]\!] = \lambda x : |x| = 1 \,.\, x$ singular

b. $[\![\text{PL}]\!] = \lambda x \,.\, x$ number-neutral

We will see that epistemic uncertainty is only one of multiple environments which reveal an asymmetry in the complexity of meaning between ϕ-featural values in the same category. In the following sections, two more contexts will be used to diagnose semantic defaults: contexts of mixed quantification and contexts of arbitrarity. We will also reveal meaning asymmetries for the category of person as well, showing that third person is the person-neutral value. This will be the task of the feature system proposed in §4.1.2: while the two notions of markedness converge for person (third person is both semantically and morphologically unmarked), they diverge for number (plural is semantically unmarked but morphologically marked). In §4.2.2, we are forced to posit a similar divergence for the inclusive.

Divergence is not necessarily deleterious. Precisely this divergence has been the object of much literature on linguistic markedness, with authors trying to relate the two notions (e.g. Bale, Gagnon & Khanjian 2011; Alexiadou 2019). We might expect complexity in forms and meanings to co-vary with each other: an increase in complexity of a form should correlate with an increase in complexity of its meaning. In other words, one might expect morphological and semantic markednesses to converge.

However, pre-theoretically, there is no mechanism that mandates their convergence[5]. In fact, a widely-assumed model of the grammatical architecture, Chomsky & Lasnik's (1977) inverted Y-model, explicitly assumes the lack of a direct link between the phonological representations (Phonological Form, PF) and semantic representations (at Logical Form, LF). Semanticists frequently propose logical operators active at LF but which are covert at PF; and in fact, resorting to lengthy circumlocutions is a frequent politeness strategy, a strategy that requires more spellout at PF. Hence, readers merely need to keep in mind that the privative features proposed in §4.1.2 are functions which operating at the level of LF, and their spell-out at PF is an independent matter.

Putting the issue of convergence aside, I will argue throughout that the relevant notion of markedness relevant for the Taboo of Directness is that of semantic markedness only. It is the task of pragmatics to determine which forms are suitably "indirect", a notion which is only sensical within the realm of meanings. After all, there is no sense in which spellouts can be "indirect" or not; hence, morphological/PF markedness will play no role in the analysis.

Semantic markedness can be cashed out in several ways. Bale, Gagnon & Khanjian (2011) model it via subset relations, where the unmarked pair member's denotation is a subset of the marked counterpart's, while Sauerland (2003) and Bobaljik & Zocca (2010) assume that unmarked pair members feature an absence of presuppositional restrictions. Although these are equivalent for my purposes, I will later choose to model semantic markedness with the latter, so that presuppositional strength is the factor being evaluated.

3.2.2 Defaults in the category of person: When does a third person not quack like a third person?

Epistemic uncertainty

Sauerland (2008: 67) proposes a test for semantic markedness based on epistemic status,

[5] Bale, Gagnon & Khanjian (2011: 206) actually refer to this possibility of convergence as the "great hope".

where the "marked member entails that the speaker is certain that some property holds, while the unmarked member only entails that the speaker takes it to be possible... To use the marked member of a morpheme pair, the speaker must be certain that its presupposition is satisfied. Otherwise the unmarked member must be used".

In the scenario in (8), a group consisting of the speaker Andreas, his addressee Bobcat, and their mutual friend Caraway are speculating about the results of a lottery. If Andreas does not know who the winner will turn out to be, the use of the third person is obligatory, as in (8c), even though the speaker Andreas or his addressee Bobcat could turn out to be the winners. In contrast, by using first or second person in (8a/b), there is no accompanying inference of epistemic uncertainty, so that Andreas must know the lottery results in advance. (The lottery context is taken from Sauerland 2008, but the test sentences have been altered.)

(8) *Context: Andreas, Bobcat, and Caraway have just bought lottery tickets. Nobody knows who will win. Andreas says to Bobcat:*

a. #Whoever wins the lottery will be so lucky... it will be **my** ticket outta this sorry-ass town.

b. #Whoever wins the lottery will be so lucky... it will be **your** ticket outta this sorry-ass town.

c. Whoever wins the lottery will be so lucky... it will be **his** ticket outta this sorry-ass town.

(9) is another example showing the same result. The context is that a group of boys have behaved disgracefully at camp, and the camp counselor responsible for them must be informed. Imagine three male camp counselors in a mixed-person group (containing the speaker Saul, his addressee Adam, and a third person Thatcher) is discussing this scenario with each other. The counselors are new on the job, and have not been assigned to specific campers, so they do not know which counselor amongst themselves is the one responsible for the disgraced boys. In a scenario where Saul is agnostic about whether it is himself, Adam, or Thatcher who will be responsible, he must use the third person form *him* to refer to the counselor eventually given this responsibility, as in (9a). In contrast, if the first or second person pronoun is used for this purpose as in (9b), there is no epistemic uncertainty involved, and the use of *whoever* becomes infelicitous.

(9) a. (Saul to Adam:) Whoever is responsible must be informed. The boys must call **him** right away.

b. (Saul to Adam:) #Whoever is responsible must be informed. The boys must call {**me** / **you**} right away.

Thus, in contexts that range over of several grammatical persons (those that feature bound variables, or those contexts without specific interlocutors or addressees), we see that the third person pronouns obligatorily appear in English.

Contexts of arbitrarity

In contexts of arbitrarity, impersonal forms are required, whether they have a dedicated ϕ-form (in English, *one*) or not (in English, *you, we*). All are equally felicitous options for e.g. prescriptions, as (10) shows.

(10) a. **One** should not wear sandals on stormy days.

b. **You** should not wear sandals on stormy days.

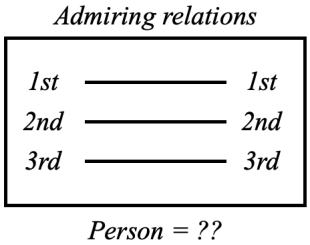
c. **We** should not wear sandals on stormy days.

Despite their surface interchangeability, there are subtle differences between each of them. In §3.2.4 we will discuss these differences, and explain the shapes of (non-)dedicated impersonals later in terms of being a default in the category of number and person, showing that plural and non-first persons are default. Until then, we will focus on the comparison between first, second, and third persons.

Mixed quantification

In contexts of a quantifier ranging over a mix of distinct grammatical persons, the bound variable that appears is in third person form. Imagine (11) was uttered in a scenario with a mixed-person group of the male speaker, the male addressee, and a male third party. In a state of affairs where each person admires themself and no others, the reflexive must be in third person form *himself*. (Note the stipulation of an all-male group fixes the gender of the third person form and avoids the notionally singular, gender-neutral pronoun *themself* as a possibility here. Without fixing the gender, *Everyone admires themself* would also be a valid substitute in this scenario.) This shows that *himself*'s range of reference may include the speaker and/or addressee of the utterance. Compare that to (11b-c) with first person *myself* and second person *yourself*, which cannot range over all grammatical persons. (11b) is only felicitous in a scenario where all three persons admire only the speaker; (11c) in a scenario where all three persons admire only the addressee.

(11) a. Everyone$_i$ admires **himself**$_i$.
b. *Everyone$_i$ admires **myself**$_i$.
c. *Everyone$_i$ admires **yourself**$_i$. (adapted from Schlenker 2005: 4)



(12) makes a similar argument. Again, when uttered in a scenario with a mixed-person group of the speaker, the addressee, and a third party, the anaphoric pronoun can only be in third person form (*his/her*).

In contrast, using the first or second person forms (*my, your*) is incompatible with mixed-person reference. In (12) below, only one mother is being called: either the speaker's mother in the case of (12a), or the addressee's mother in the case of (12b).

(12) a. [Every one of us]$_i$ has to call **my**$_i$ mother.
b. [Every one of us]$_i$ has to call **your**$_i$ mother.

Grammatical person and imposters

As mentioned, imposters are third-person expressions which can take on first or second person reference (e.g. Collins & Postal 2012, Podobryaev 2017). In (13), the third person DPs *the judge* and *Father* take on first person reference. In (14), the third person DPs *the reviewer* and *Grandmother* take on second person reference. Despite their notional reference, note that they trigger third person verbal agreement throughout, showing that imposters are indeed grammatically third person.

(13) a. (at an awards ceremony) **The judge** is (=I am) now announcing the prize.
b. (to son) **Father** is (=I am) coming to get you.

(14) a. (in a reply to one's reviewer) **The reviewer** is (=You are) being overly anal.
b. (at family gathering) I can understand why **Grandmother** is (=you are) upset; I'm sorry.

In rhetoric, illeism is the practice of referring to oneself in the third person, which is deployed for a range of stylistic purposes. (15a) is taken from a Seinfeld episode where Elaine dates a man, Jimmy, whose habit of speaking about himself in the third person indicates his arrogance, and causes a wide array of comic confusion. Jimmy's illeism turns out to be infectious; George acquires this manner of speaking at the end of the episode (15b).

(15) Seinfeld S06E18, "The Jimmy"
a. Jimmy: **Jimmy's** gonna get you...! Hands off **Jimmy**! Don't touch **Jimmy**! (=I'm gonna get you...! Hands off me! Don't touch me!)
b. George: **George** likes spicy chicken.
Jerry: What's that?
George: **I** like spicy chicken.
Jerry: No, no, you said, "George likes spicy chicken."
George: No, I didn't!
Elaine: Yes, you did. You said, "George likes spicy chicken."
Jerry: You're turning into Jimmy!
George (irritated): **George** is getting upset!!

An interesting example of illeism comes from Sméagol/Gollum, a character from *Lord of the Rings*. (16a-c) features the ungrammatical strings *we + V-s* in his idiolect. The subject pronoun appears in the first person plural, which is a trope pervasive in his idiolect to indicate his split personality. However, verbal agreement is definitively third person singular *-s*, so that illeism is expressed via verbal agreement instead of via the subject pronoun[6].

(16) (Sméagol/Gollum:)
a. **We wants** it, **we needs** it. Must have the precious. Tolkien The Two Towers
b. My birthday-present! That's what **we wants** now, yes; **we wants** it! My birthday-present! It came to me on my birthday, my precious.'
c. Goblinses down there. Lots of goblinses. **We smells** them. Ssss! Tolkien The Hobbit, or There and Back Again

In addition to indicating self-grandiosity, illeism is also frequently deployed to signal childlike behavior or simple-mindedness. The star of the children's show *Sesame Street*, Elmo, always refers to himself in the third person on air (17a). Meanwhile, the simple-minded brute Hulk, from *The Incredible Hulk*, frequently refers to himself in the third person (17b).

(17) a. (Elmo:) **Elmo** has (=I have) a question.
b. (Hulk:) **Hulk** (=I) smash. **Hulk** (=I) smash EVERYTHING!

[6]Additionally, Sméagol/Gollum does not refer to himself in the second person either. In (ib), Sméagol/Gollum uses third person *it* instead of *you* to refer to his addressee Bilbo.

(i) a. Bilbo: What have I got in my pocket?
b. Sméagol/Gollum: It must give us three guesseses [sic].

Tellingly, there is no practice of referring to oneself in the second person, and indeed, imposters cannot take on the shape of first or second persons. There is no context, stylistic or otherwise, where substituting *I, we, you* would be licit for non-first or non-second reference. In (18a), a conference participant is waiting for the others to announce a prize; *we/you* cannot be used in place of a third person pronoun. In (18b), an errant grandchild is apologizing to her grandmother; *me/you* cannot be used in place of *Grandmother*.

(18) a. (at a conference) #**We/You** are (=She is) now announcing the prize.
b. (at family gathering) #I can understand why **I am/You are** (=Grandmother is) upset; I'm sorry.

A naturalistic example comes from a handwritten thank-you note from a mother-daughter business, where the mother-daughter team thanks a customer for their patronage. The instances of *we* used may be either a choric first person or an exclusive plural, but crucially, when referring to each of the two members of the team, third person *she* is used for both conjuncts (instead of alternating *she* with *I* between the conjuncts). This shows that third person *she* imposters for first person *I* at least once.

(19) We're a Mum (**she** sews) and daughter (**she** climbs) team, and we've had so much fun designing and creating this chalk bag just for you.

3.2.2.1 Interim summary for person: Third person does not quack like a third person when it is person-neutral

From §3.2.2, we saw that it is third person, not first/second person, which must surface in default-hugging contexts: in contexts without specific interlocutors, in mixed quantifications, and in epistemically uncertain contexts. Furthermore, third-for-first and third-for-second both seem to be productive stylistic mismatches, while mismatches in the opposite directions are unattested. We therefore conclude that within first, second, and third persons, third person is the semantic default, being able to take on person-neutral reference. This conclusion is also compatible with previous literature on the topic (Sauerland 2008). Now, we move on to diagnosing defaulthood in the category of number.

3.2.3 Defaults in the category of number: When does a plural not quack like a plural?

Naturally, number has been discussed in conjunction with person in the literature on semantic markedness. In singular-plural number systems, plural is said to be semantically unmarked (Krifka 1989; Croft 2003; Sauerland 2003, 2008b; Sauerland, Anderssen & Yatsushiro 2005; Spector 2007; Zweig 2009; Bale, Gagnon & Khanjian 2011[7]; *a.o.*).

It is important to note that this body of existing literature has only focused on number marking in the nominal domain, as §3.2.3.1 will show. In §3.2.3.2, I provide evidence that the defaulthood of plural also extends to the *pronominal* domain, particularly pertinent here because the typology is one of polite pronouns.

[7]Bale, Gagnon & Khanjian (2011) originally concluded that in Turkish and Western Armenian, it is the singular which is semantically unmarked, not the plural. However, Yatsushiro, Sauerland, and Alexiadou (2017) later disputes this conclusion. This dispute in the literature is noted here as a disclaimer, necessary since the project to investigate semantic markedness cross-linguistically is ongoing.

3.2.3.1 Nominal number

Epistemic uncertainty

The plural surfaces when speaker is uncertain about the cardinality that holds of a referent. Consider the monologue in (20), where the speaker is contemplating pet hamsters past and present.

(20) How many **hamsters** do I own?

a. Just **one hamster**, I think.

b. ... Last month the breeder promised an increase in numbers.
(sarcastically) I've seen an increase all right, an increase of **-2 hamsters**.

c. ... Last year, I owned only an average of **0.5 hamsters** per month.

d. ... (discovering escape of last remaining hamster) Oh no, actually, **0 hamsters**.

(20a) contains a singular NP (*one hamster*). Yet, it is still a felicitous answer to the question containing a plural noun (*hamsters*), showing that the plural form can be felicitously used to inquire about individuals. (20b-d) are also possible answers: even though the cardinalities of hamsters are neither pluralities nor individuals (-2, 0.5, 0), plural marking is obligatory. This shows that plural is compatible with a wider range of uses, while singular marking is only compatible with a cardinality of 1 as in (20a).

When it is known that a referent is a singular, the plural form may not be used, only the singular. The context of (21) is one where the title of Mr. America is given to the contestant with the highest points in the Speedo competition[8].

(21) {**The winner** / #**The winners**} will be insanely rich.

In contrast, when it is unclear that there will only be one winner or multiple, the plural must be used. The context of (22) is one where a title of Mr. America is given to any contestant with scores above the 95th percentile.

(22) {#**The winner** / **The winners**} will be insanely rich.

In similar contexts of discovery, it is again the plural that must be used. A strapping young ad-man, freshly employed at NASA, must convey the possibility of endless discovery in space. In his copy, he must use (23a) with the plural—NASA has the task of discovering as many new galaxies as possible, but (23a) is also felicitous if NASA discovers exactly one new galaxy and no further. In contrast, if he uses (23b) with the singular, this implies NASA's time is limited (that after discovering one galaxy, the agency will shut down) or that there is only one new galaxy that remains to be discovered. (Since galaxies do not naturally come in singularities or in groups, the confounding factor of expected cardinalities is also removed here.)

(23) (NASA ad:)

a. The **galaxies** that we discover tomorrow might be teeming with life.

b. The **galaxy** that we discover tomorrow might be teeming with life.

Consider also the billion-dollar startup Colossal Biosciences, which has promised to genetically

[8] An alternative to the Mr. America scenario is the Powerball lottery, but it is a common practice for multiple individuals to pool their money together to buy a single lottery ticket, destroying the intended setup where there can only be a single unique winner. Hence Mr. America.

resurrect the woolly mammoth by 2027 (along with several other extinct species like the dodo, the Tasmanian tiger, and the northern white rhinoceros). While it would already be a huge feat of genetic engineering to de-extinctify a single woolly mammoth individual, company copy would likely use (24a) with the plural over (24b) with the singular. This is because the latter implies that the company has the narrow ambition of resurrecting exactly one individual and no more, a stance bad for business.

(24) (potential Colossal Biosciences ad:)
  a. The **mammoths** that we recreate will be a huge win for biogenetics.
  b. The **mammoth** that we recreate will be a huge win for biogenetics.

Hence, we see that plural is number-agnostic, not singular.

Contexts of arbitrarity?

[9]This particular diagnostic is unfruitful for assessing default number marking on nouns, because, as we will show, nouns are accompanied by an expected cardinality which is determined by the context. It is very difficult to escape the expected cardinality of a noun, as we show below.

At first glance, English prefers the plural in utterances without specific cardinalities, such as public prescriptions. We can easily imagine plural (25a) written outside a family clinic, where the clinicians very well know that families may come with one or more strollers depending on the number of children they have. Similarly, plural (25b) is expected in at an automated gas station, where one may insert one or multiple bills, depending on the size of the bill's denomination and total price.

(25) a. Park your {**?stroller** / **strollers**} here.
  b. Insert your {**?bill** / **bills**} here.

However, we cannot take the above too hastily to be evidence for plural as a default, as pragmatic naturalness often makes a difference for such judgments. In (25b), for instance, the plural preference may come from the fact that one rarely has a single bill that fits the bill just right. In front of an ATM, it would instead be odd to see the sign with a plural (*?Insert your credit cards here*), as only one credit card is required at ATMs, even if the user has multiple credit cards. Thus, an expectation with regards to an object's cardinality cannot be removed in public signs, which necessarily preside over a specific object.

Dramatically, the same noun can be preferred with the singular or the plural depending on context. Bale, Gagnon & Khanjian (2011: 203) give an example with the noun *umbrella*, which I adapt here. Imagine that Alexis is calling her friend Borax, and that neither Alexis nor Borax routinely use their dishwashers as a drying rack for wet umbrellas. If Borax is asked (26), he would hesitate to answer affirmatively even when there is only one umbrella in his dishwasher, as umbrellas are not naturally expected in dishwashers.

(26) (Alexis to Borax:) Do you have **umbrellas** in your dishwasher?

However, if Alexis is asking a shopowner, Sappy, about the availability of umbrellas for sale as in (27), this hesitation disappears. Since umbrellas are naturally expected in stores, Sappy can answer affirmatively even if there is only one umbrella left in the store.

(27) (Alexis to Sappy:) Do you have **umbrellas** in your store today?

[9]Many thanks to Nina Haslinger for her extensive guidance in several of these issues.

The same facts about pragmatic naturalness can be shown with the noun *pool*, where it is understood that (indoor) pools are more commonly understood to be found in sport complexes instead of in private residences. Even if there turns out to be only one pool, (28b) is more natural than (28a). (Note that the oddness of (28a) cannot be only attributed to the fact that most people do not typically have indoor pools in their houses: while it is indeed an atypically opulent possession, it is not completely unattainable. The sentence with the singular *Beware, the pool in this house is overly chlorinated* is perfectly natural.)

(28) a. #Beware, the **pools** in this house are overly chlorinated... Actually, I can smell **it** already.
b. Beware, the **pools** in this sports complex are overly chlorinated... Actually, I can smell **it** already.

For objects that naturally come in singulars, like certain body parts and certain kin, the plural form is highly marked. Hence, (29a-c) with *nose, bellybutton* and *mother* are odd. Note that (29c)'s oddness is not ameliorated by much even if we understand that Nigel was adopted by two women.

(29) a. #My **noses** get red in the winter from the cold.
b. #My **bellybuttons** can accumulate a lot of dirt.
c. #Nigel's **mothers** pick him up from kindergarten every day.

The singular counterpart is much more felicitous in these cases:

(30) a. My **nose** gets red in the winter from the cold.
b. My **bellybutton** can accumulate a lot of dirt.
c. Nigel's **mother** picks him up from kindergarten every day.

Conversely, when it is understood that the expected cardinality is more than one, than the use of the singular becomes highly marked. In an airport security context where one must place four fingers simultaneously on the scanner for fingerprinting, (31a) with the singular is infelicitous; (31b) with the plural must be used instead.

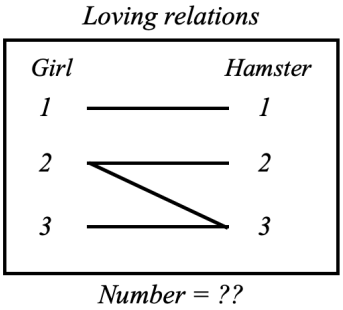
(31) a. #Please place your **finger** on the scanner.
b. Please place your **fingers** on the scanner.

Hence, evidence the nominal domain does not conclusively indicate that the plural is number-agnostic. What is relevant for determining the grammatical number that appears in the above cases is our expectations for an object with regards to its cardinality. In contexts where the singularity is the expected cardinality, the plural cannot be used to quantify over these singularities, and vice versa. Hence, we proclaim that this diagnostic is inapplicable to assess defaulthood of number marking on nouns, and move on to the diagnostic of mixed quantification.

Mixed quantification

Consider also the shape of bound variables in quantificational contexts. In scenarios involving mixed numbers, we predict that the semantically unmarked number value surfaces. Consider (32) with the universal quantifier *every*:

(32) a. Every girl loves **her hamster**.
b. Every girl loves **her hamsters**.



(32b) is compatible with a scenario with only 1-to-1 relations: each girl owns exactly one hamster. However, it is also compatible with a mixed scenario depicted in the diagram, where some girls own exactly one hamster (singular relations) and other girls own multiple hamsters (plural relations). In such a mixed scenario, plural marking is obligatory, showing that plural is less marked. Meanwhile, (32a) is only compatible with a scenario containing only singular relations, where each girl owns exactly one hamster.

Non-canonical ranges of reference

The number-neutrality of the plural can also be appreciated with sentences with non-canonical ranges of reference, a type of mixed quantification resulting from the use of conjunctions or disjunctions.

We can appreciate the same fact with numerical ranges containing 1. (33) is a disclaimer from a pharmaceutical company, which appears on the product packaging of their morning-after pill. Even though 1 micro-milligram is contained within the range, the plural form *micro-milligrams* must be used.

(33) Please be advised that this product contains 1–10 {***micro-milligram** / **micro-milligrams**} of testosterone, but this will not affect its overall efficacy.

In (34), numerical range is expressed via disjunction. Imagine (34) where an exasperated parent is speaking to a child with a particular reptile hankering; only the plural is allowed.

(34) Fine. But whether you get one or three {***reptile** / **reptiles**}, you're going to be the one responsible for all the vet fees.

The obligatory use of plural in (33)–(34) might be merely due to a closest conjunct or disjunct agreement effect; however, we can guard against this possibility by setting to 1 the conjunct or disjunct closest to the noun. The preference for plural over singular persists even in these cases, as in (35)-(36).

(35) Please be advised that this product contains 0.3–1 {***micro-milligram** / **micro-milligrams**} of testosterone, but this will not affect its overall efficacy.

(36) Whether you get three or one {***reptile** / **reptiles**}, you're going to be the one responsible for all the vet fees.

Nominal number and imposters

There is also a case to be made for over-pluralization being used for stylistic purposes. (37a) presents a quote taken from a boss-level character, Legion, in the Nintendo video game *Fire Emblem: New Mystery of the Emblem*, showing that Legion's idiolect involves over-pluralization,

so that the noun is pluralized even though its determiner is singular. (37b-c) present quotes from Sal, a blue-collar worker in the *Futurama* series who is portrayed as lazy, poor, and incompetent. His idiolect features over-pluralization as a marker of his low social class.

(37) a. (Legion:) We may have lost **this battles**.
b. (Sal:) Has yous ever seens **Bigfeet**?
c. (Sal:) Don't worries, I've gots an idea, I mean, **an ideas**.

(38) is taken from *A Bit of Fry and Laurie*, where a man walks into a shoe shop in search of shoes, but discovers that it is actually a front for a brothel. The shopkeeper breaks this news to him, but tacks on the plural morpheme randomly, presumably to avoid the negative connotations associated with the word *brothel*. The effect of the plural is that it gives the shopkeeper an air of childish naïveté, suitable for keeping up a front.

(38) I dislike the word 'brothel', Mr. Jowett. I prefer the word 'brothels'. Yes, this is **a brothels**. (S03E03, *Shoe Shop*)

In contrast, over-singularization is never used for stylistic purposes. I do not know of any e.g. fictional character who uses singular forms to refer to plural entities, as in (39).

(39) Gosh! All those **fighter** have run away!

3.2.3.2 Pronominal number

The above tests show that plural is semantically unmarked relative to the singular. However, since the typology in Chapter 2 is one of polite *pronouns*, we must make sure that the results can be replicated from the nominal domain to the pronominal domain. This is especially because much of the theoretical and experimental literature on the semantic markedness of number has only considered the markedness profile of nominal plurals to the exclusion of pronominal plurals. For instance, Sauerland, Anderssen & Yatsushiro's (2005) experimental stimuli did not include pronominal plurals, only nominal plurals. Thus, the claim that plural is a semantic default should be made with caution. To ensure that nominal number and pronominal number can be placed on equal analytical footing, we show that the results above can be replicated for pronominal number. It is this task to which we turn now.

Epistemic uncertainty

When it is known that the referent is a singularity, the plural pronoun may not be used, only the singular (40). In case of uncertainty, the plural pronoun may be used, and not the singular (41). We replicate the Mr. America context from above, this time varying the number marking on pronouns instead of on nouns.

(40) *Context: The title of Mr. America is given to the contestant with the highest points in the Speedo competition. Everybody knows there will only be one Mr. America. A contestant says:*
a. The winner will be so lucky. It could be **me**.
b. #The winner will be so lucky. It could be **us**.

(41) *Context: A title of Mr. America is given to any contestant with scores above the 95th percentile. Nobody knows how many winners there will eventually be; it could one or multiple. A contestant says:*

a. #The winner will be so lucky. It could be **me**.
b. The winner will be so lucky. It could be **us**.

Again, in the disjunctive scenario where an exasperated parent is speaking with a child with a particular reptile hankering; only the plural pronoun is allowed.

(42) Fine. But whether you get one, or three, you're going to be the one responsible for {#**him** / **them**}!

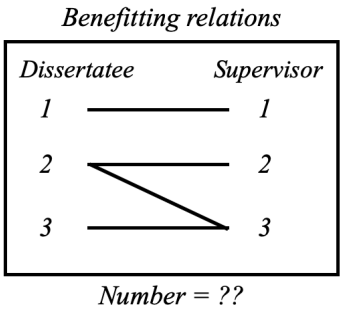
Mixed quantification

Consider a variation of (11) above. As before, imagine that (43) was uttered in a scenario with a mixed-person group of the male speaker, the male addressee, and a male third party. The third person plural form *themselves* in (43a) is judged more natural[10] than with the singular *himself* in (43b).

(43) a. Everyone$_i$ admires **themselves**$_i$.
b. Everyone$_i$ admires **himself**$_i$.

Imagine another scenario with mixed quantification, as in (44), where a group of dissertating students are discussing a set of particularly great supervisors. In a scenario where there is a mix of 1-to-1 benefitting relations *and* non-1-to-1 benefitting relations, as schematized, only the plural is suitable as in (44a). To use (44b) with the singular would be inappropriate, as it wrongly assumes there only exist 1-to-1 benefitting relations.

(44) *Context: Dissertating students discussing a set of particularly great supervisors. Some dissertatees work with 1 of these, while dissertatees work with ≥1 of these.*
a. Every dissertatee benefits from **them**.
b. Every dissertatee benefits from #**her**.



Pronominal number and imposters

In rhetoric, nosism is the practice of referring to oneself in the plural. One example of this is 'editorial *we*', as in (45a). There are also examples of fictional characters using the plural pronoun *we* instead of *I*, to indicate self-grandiosity. In (45b), Legion is a boss-level battle character in Fire Emblem (a Nintendo video game franchise) refers to himself as in the plural.

(45) a. (from the desk of a publisher:) **We** are (=I am) writing to inform you that your short story will not be included in our Fall 2026 issue.

[10]It should be noted that other third person plural alternatives are also acceptable here, such as *themself* (due to the rise of singular *they*), and *theirself, theirselves* in some English dialects.

b. (Legion:) Are **we** real or are **we** false? Try **us** and see. (=Am I real or am I false? Try me and see.)

Tellingly, there is no practice of referring to groups in the singular, as in (46).

(46) a. (from the desk of Prince Charles and Princess Diana:) ***I** am (=We are) announcing our divorce effective immediately.
b. (a band of low-level battle characters:) ***I** am (=We are) small but many!

3.2.3.3 Interim summary for number: Plural does not quack like a plural when it is number-neutral

We see that it is plural, not singular, which must surface in contexts triggering the appearance of a default: in contexts without expectations of cardinality, in mixed quantifications, in epistemically uncertain contexts, and in numerical ranges containing multiple grammatical numbers. This holds in both the nominal and pronominal domains. This suggests that the denotation of plural contains that of the singular, and that plurals are not restricted to exclusive interpretations (strictly *more than one*). Furthermore, plural-for-singular seems to be a widely attested stylistic mismatch, while the mismatch in the opposite direction is unattested. We hence conclude that plural is the semantically unmarked/default number, being the number value suitable for number-neutral interpretations.

3.2.4 Shapes of arbitrary reference: Many quacks given

In this section, we consider the diagnostic of arbitrarity in the pronominal domain by examining the ϕ-featural shapes of impersonal pronouns. As was flagged above in (10), the (English) picture is complicated by the fact that impersonal reference can be achieved via several ϕ-forms, each with different properties. We establish two hallmarks of impersonals in §3.2.4.1 and distinguish between routes to impersonality (following Malamud 2006; Zobel 2014; *a.o.*) in §3.2.4.2.

3.2.4.1 Hallmarks of impersonality

Impersonals with dedicated forms feature two distinct semantic properties. The first is an humanness requirement[11]. Thus, in (47), English *one* and German *man* cannot refer to a newspaper or a sign that are impossible to decipher, but necessarily to a person who is impossible to understand. Hence, (47) is acceptable if the speaker is expressing frustration about the locals speaking an unfamiliar dialect, and not acceptable if the speaker is expressing frustration about written material in an unfamiliar script/typeface/scrawl.

(47) a. **One** is incomprehensible.
b. ***Man*** *ist* *unverständlich.*
IMPERS COP.3SG incomprehensible

Positing an animacy requirement instead of a humanness requirement would be too weak. As (48) shows, English *one* and German *man* cannot refer to e.g. a baboon, but necessarily to a person.

(48) a. **One** hoards food during shortages.

[11]Thank you to Daniel Harbour for pointing this out, and to Kai von Fintel for providing a suitable example.

b. ***Man*** *hortet* *Nahrungsmittel bei Engpässen.*
IMPERS hoard.3SG food by shortages

The second is that these pronouns are number-agnostic, making the label "impersonal" partially a misnomer. This is despite taking 3SG verbal agreement in European languages, as the German examples show clearly in (47b) and (48b). There is good reason to think that 3SG verbal agreement is just a morphological default, as these impersonal pronouns can refer to one or multiple human individuals, as the felicity of the continuations in (49) shows for English. To my knowledge, no language possesses dedicated impersonal pronouns that are morphologically differentiated for number itself.

(49) One speaks German here, although {he is / they are} speaking in a dialect of German I don't understand.

Furthermore, it cannot be the case that the impersonal can only refer to a singular third person: (50) shows that the impersonal is person-neutral as well, being able to take on first and second person reference.

(50) a. One hoards food during shortages, but {I / we} must be considerate of those who are elderly or pregnant.
b. (to one addressee or multiple addressees): One hoards food during shortages, but {you} must be considerate of those who are elderly or pregnant.

To summarize, the semantic properties of dedicated impersonals are that i) their referent is necessarily human and ii) they are number- and person-neutral.

3.2.4.2 Routes to impersonality

Moderator: *Do you feel it's unfair that we look to you for answers on how to rebuild trust in science, or do you accept it as part of your fate?*

Fauci [visibly eepy]: *It depends on what you mean by 'we'!*

— WBUR festival, May 31st 2025

In this section, we will distinguish two routes to impersonality: distributive quantification and plural predication. The former route is taken by morphological singulars while the latter is taken by morphological plurals. We consider both, but only the latter will be relevant for our purposes of semantic defaulthood, as only the latter represents a form of underspecification (for number).

English is a language which takes both routes to impersonality. It features a dedicated impersonal pronoun *one*, which is somewhat restricted to higher registers, but also recruits the personal pronouns *you* and *we* for impersonality. Consider the general social prescriptions in (51)-(52), which are equally felicitous with any of the three options (although the use of *one* may be dispreferred due to its additional register restriction).

(51) a. **One** shouldn't dismiss midwifery as a profession!
b. **You** shouldn't dismiss midwifery as a profession!
c. **We** don't dismiss midwifery as a profession!

(52) a. **One** doesn't mention Ms. Voldemort here.

b. **You** don't mention Ms. Voldemort here.
c. **We** don't mention Ms. Voldemort here.

One naturalistic example comes from *The Crown* (S06E01), in the Prince of Wales' speech for Camilla Parker-Bowle's 50th birthday (53a). The instances of *one* are interchangeable with *you/we* (53b, c) (note that the Prince himself switches to *we* in the second sentence anyway).

(53) a. As **one** gets older, **one** starts to realize what **one** really values in life: **one's** health, **one's** family, and **one's** friends... The people who enrich **our** lives and who **we** can't bear to live without.
b. As **you** get older, **you** start to realize what **you** really value in life: **your** health, **your** family, and **your** friends... The people who enrich **your** lives and who **you** can't bear to live without.
c. As **we** get older, **we** start to realize what **we** really value in life: **our** health, **our** family, and **our** friends...

So, what differentiates these distinct impersonal pronouns from each other? Following Kitagawa & Lehrer (1990), we distinguish between impersonal reference achieved via either distributive quantification or plural predication.

The first route to impersonality, distributive quantification, ascribes some property to every individual in the relevant domain. These are *one* and *you* in English. For instance, Malamud (2006) argues that the arbitrariness of English impersonal *you* stems from its origin as a variable in its denotation, which later is associated with a discourse function of non-participation. Thus, it is not born associated with second person ϕ-features (she leaves its second person spellout aside). Meanwhile, the second route to impersonality, plural predication, is more on par with existential quantification: it picks out a plurality and ascribes a property to that plurality. No specific individuals are identified or identifiable. This is *we* in English (also English *they*, which we cover later).

One test to tease apart these two routes is whether it is acceptable to exclude the speaker from the general group[12]. For social prescriptions made with plural impersonals like *we*, it is impossible to for the speaker to exclude themselves from the prescription, so that the speaker of (54a) must be a trailblazing fashionista herself. However, exclusion is possible for those made with distributive impersonals like *you*, so that the speaker of (54b) can either be a trailblazing fashionista or not.

(54) a. As trailblazing fashionistas, **we** must make **our** selections carefully and appropriately. **We** would not wear formal clothes at the beach or bathing suits in church.
b. As trailblazing fashionistas, **you** must make **your** selections carefully and appropriately. **You** would not wear formal clothes at the beach or bathing suits in church.

Similarly, (55) with the plural impersonal *us* is only felicitous if reported by speaker of the same nationality as the interlocutor; in this case, American, as it was printed by the newspaper *USA Today* in 1986.

(55) Nationwide only 7.8% of **us** are without a telephone at home. (Kitagawa & Lehrer 1990: 745)

The point can be made most clearly when we have groups necessarily differentiated by biology (i.e. men and women). Thus, a general prescription like (56a) with the plural impersonal *we*

[12]Thank you to Sarah Zobel for bringing this diagnostic to my attention.

can only be uttered by a fellow Y-chromosome holder. (56b) with the distributive impersonal *you* can be uttered by a fellow Y-chromosome holder or not. Thus, exclusion is only possible with (56b).

(56) a. Y-chromosome holders should band together and take charge of **our** longevity.
b. Y-chromosome holders should band together and take charge of **your** longevity.

Some dedicated impersonal pronouns may take the plural route to impersonality, as they disallow exclusion of the speaker from their reference. This patterns away from English *one*, such as the French impersonal *on* and the German impersonal *man*. The sentences in (57) (Malamud 2006: 9) are compatible with a scenario where the speaker is talking about family customs which she herself follows.

(57) a. *In meiner Familie spricht* ***man*** *höflich miteinander.*
In my family talks **IMPERS** politely with.one.another
'In my family, we speak politely with each other. German
b. *Dans ma famille,* ***on*** *se parle entre soi.*
In my family **IMPERS** MIDDLE talks between oneself
'In my family, we talk to each other.' French

Lastly, the ϕ-featural combination of 3PL may take yet another route to impersonality. Namely, as it combines the semantically default values from person (third person) and plural (number), it is also a semantic default default itself. In English, 3PL *they* can also be used for impersonal reference. For instance, the expression *(As) They say...* prefaces pieces of general wisdom or popular convention, instead of referring to a specific third party or a specific group of third parties (58a-c). (58d) is the title of a *Washington Post* podcast episode discussing the administration's inscrutable immigration policy in March 2025, where *they* refers to a group of unspecified decision makers within the administration.

(58) a. **They** say that babies should not bathe alone.
b. **They** say to only sleep 7 hours a day.
c. (of March Madness) As **they** say, stay tuned. (source)
d. Just come in legally, **they** said (source)

Importantly, the reference of 3PL *they* is both person-neutral and number-neutral, merely referring to the relevant people, vaguely defined. In (58a), for example, the entity spreading the wisdom to not let babies bathe alone might be an individual (e.g. a celebrity pediatrician on TikTok) or a group (e.g. several pediatricians)[13].

In Italian and Russian, 3PL is the ϕ-feature that appears in arbitrary contexts with antecedentless pronouns. Thus, the sentences in (59) do not indicate a specific plurality of referents, but rather to a vague general population: that is, those unspecified individuals that speak English by dint of being in America, whether they be American citizens, immigrants, international students, confused tourists, etc.

(59) a. *V Amerike* ***govor'at*** *po-anglijski.*
In America speak**.3PL** in-English
b. *In America* ***parlano*** *inglese.*
In America speak**.3PL** English

[13] Also noteworthy is the fact that impersonal *one* in English can only be bound by 3PL *they*, as in *As trailblazing fashionistas,* ***one*** *must make* ***their/*your/*our*** *selections carefully and appropriately.* I do not know what to make of this, and only note it here.

'They (generic) speak English in America.' (Malamud 2006: 2-3)

To summarize, several pro-forms appear in contexts of arbitrarity, falling into at least three groups. First are those forms which achieve impersonal reference via distributive quantification (*you, one*) which are morphologically singular for that reason. Second are those forms which achieve it via plural predication (*we, they*) which are morphologically plural for that reason. Third are those forms which achieve impersonal reference by dint of embodying a combination of person-neutral person and number-neutral number (*they*).

This discussion merely acknowledges the existence of several routes to impersonality possible. Note that only the second and third groups of impersonals fit the established trend of plural and/or third person being semantic defaults. This is important for our discussion, because it fits the conception of semantic defaults appearing in contexts of arbitrarity (*i.e.* contexts which license impersonals). Does the first group of "singular" impersonals constitute a counterexample, then? We will assume not: the existence of "singular" forms appearing in contexts of arbitrarity does not buck this trend, as their singularity is ascribed to their distributivity.

3.2.5 Summary of default feature values

Table 3.1 summarizes the results of the diagnostics as applied to ϕ-featural values in the pronominal domain. ⊗ indicates that a ϕ-featural value is not a default, while ✓ indicates that a ϕ-featural value is a default.

Which ϕ-featural value(s) may appear:
- in contexts of epistemic uncertainty;
- in contexts of mixed quantification;
- in contexts of arbitrarity;
- as imposters?

	⊗	✓
person	1, 2	3
number	SG	PL
clusivity	EXCL	INCL

Table 3.1: Semantic defaults for person, number, and clusivity

While not all diagnostics were applicable to all ϕ-featural categories, the diagnostics that were applicable for each category yielded consistent results. For person, third persons and impersonals are semantic defaults, not first/second persons. For number, whether it be nominal or pronominal, plural is the semantic default, not singular. Lastly, for clusives, inclusives are semantic defaults, not exclusives.

The exclusion of downward-entailing environments

Reader may wonder why downward-entailing environments (such as negation, and antecedents of conditionals)[14] do not feature as a diagnostic group, since they may potentially illuminate relations of logical strength between feature values for the following reason. If it is possible to switch A with B in a downward-entailing environment and preserve truth, then B has a meaning that entails that of A. This shows that A has a wider denotation than B.

[14]Questions have also been commonly used as downward-entailing environments in the literature, even though they do not have "truth conditions" in the traditional sense, but because they license NPIs, a property famously linked to downward-entailing environments by Ladusaw (1980). Thus, accepting questions as a downward-entailing environment is highly dubious. This irregularity, if one wants to call it that, has more implications for the semantics of questions, though, than for monotonicity.

Such a frequently used environment is negation: (60a) with the singular is compatible only with a scenario where Sue has exactly one pet hedgehog, while (60b) with the plural is compatible either with a scenario where Sue has exactly one pet hedgehog, or where Sue has more than one.

(60) a. Sue didn't bring **her hedgehog** to the party.
b. Sue didn't bring **her hedgehogs** to the party.

Such environments are not good test environments for semantic defaults for a reason already outlined above, which is that nouns carry with them expectations of cardinality. Again, we can appreciate this by considering nouns which typically denote unique entities, such as body parts or kinship terms. In these same environments, using the plural forms are markedly odd: (61a) cannot be used to describe the typical scenario of having exactly one bellybutton, and (61b) cannot be used to anti-identify Nigel's caregivers, even if the speakers are well-acquainted with the scenario of lesbian relationships, or of having two female primary caregivers.

(61) a. I don't have **bellybuttons**.
b. Those people over there are not **my mothers**.

Downward-entailing environments are hence not as immediate a diagnostic for semantic markedness as the literature takes them to be. Our objective is to examine contexts that do not specify or supply specific feature values—namely, those contexts where *Maximize Presupposition!* is unable to apply—and downward-entailing environments like negation and questions are not such contexts, hence their exclusion as diagnostic environments.

When we attempt to remove the confounding factor of expectedness from downward-entailing environments, we arrive at the context of epistemic uncertainty—precisely one of the diagnostics proposed for semantic defaulthood above. Imagine that a football coach presides over a male cheerleading team[15]. However, the coach has received complaints regarding misogynistic behavior in the locker rooms. Compounding his disappointment is the fact that he does not know how many cheerleaders were complained about. (62) shows that, in this scenario, the plural is compatible with two scenarios: one cheerleader, or more than one cheerleaders, being penalized. (Here, we use *any of them* for the plural. This choice is due to *them* having a singular use in English, yet a reinforcement like *them all* triggers additional inferences about collectivity. It suffices for our purposes that *any of them* is not compatible with a singular interpretation of *them*.)

(62) Whoever is exhibiting misogynistic locker room behavior... I simply will not bring **any of them** on for halftime.

However, the permissiveness of the plural in (62) is not due to negation, but due to the epistemic uncertainty of the coach—exactly one of the diagnostics proposed for semantic defaulthood. (63) with the singular is simply incompatible with a scenario where the coach is uncertain about how many cheerleaders were complained about; it is only compatible with one scenario: where one cheerleader is being penalized.

(63) Whoever is exhibiting misogynistic locker room behavior... I simply will not bring **him** on for halftime.

Hence, while downward-entailing environments have been much-utilized in the literature on

[15]Despite the stark, 100% gender bias in current North American cheerleading teams, which are all-female, the very first cheerleading squad was all-male. They started cheering for Princeton in 1965, before Princeton admitted women, in 1969. This context thus harks back to a time with gender inequality of a different kind.

semantic markedness, they are not environments which necessarily trigger the appearance of semantic defaults. I define contexts which *do* trigger the appearance of semantic defaults as contexts where *Maximize Presupposition!* is unable to apply for one reason or another: contexts of epistemic uncertainty, mixed quantification, and arbitrarity are used as diagnostic environments instead.

3.3 Diagnosing semantic defaults, cross-linguistically

So far, we have established third person and plural as default ϕ-featural values only on the basis of English data. Since the recruitment asymmetries in the shapes of polite pronouns in Chapter 2 are typologically motivated, a note on the universality of semantic defaulthood is in order. Specifically, the shapes of semantic defaults established in §3.2 must be commensurate in empirical scope. This is because the impending analysis of semantic defaults will be intimately linked to the less-impending analysis of polite pronouns.

Given practical constraints, I cannot present language-specific evidence for semantic defaulthood for all >200 languages considered in Chapter 2. In principle, to formulate any language universal is an unattainable goal, given the impossibility of obtaining negative evidence for each language. However, stipulations of this type are necessary for theoretical work to progress, and making this assumption is neither controversial nor futile.

Non-controversially, most cross-linguistic work on the topic concludes that plurals are semantic defaults. A large body of work (Krifka 1989; Croft 2003; Sauerland 2003, 2008; Sauerland, Anderssen & Yatsushiro 2005; Spector 2007; Zweig 2009; Sudo 2023) establishes nominal plural as a semantic default; Yatsushiro, Sauerland & Alexiadou (2017) present further experimental evidence in 18 European languages supporting this hypothesis. The current work merely presents evidence extending this hypothesis to pronominal plurals as well. The conception of third persons as semantic defaults is also compatible with existing work on the topic (Sauerland 2008). A further contribution of this dissertation is the identification of impersonal as another semantic default in the category of person. (Owing to a dearth of relevant work, the universality of inclusives as defaults is not so clear, and this is earmarked as a novel problem in §4.2.2.)

Making this assumption of universality is also not futile, as it enables us to make concrete and testable predictions. In particular, *a semantically default value will always be fair game for the recruiting mechanism, modulo* language-specific constraints discussed later in Chapter 4.

With that said, this section presents novel data on semantic defaulthood in the ϕ-featural domain, focusing on non-European languages, and on the clusivity distinction. While it is extremely difficult to find any discussion of epistemic uncertainty and mixed quantification from descriptive grammars, it is more straightforward to find descriptions on how arbitrary reference is achieved, and imposter phenomena. This section will view these two diagnostics through a typological lens.

Two types of convergences will emerge. The first convergence is a cross-linguistic one, where the shapes of semantic defaults converge to the current English-based results. The second "bonus" convergence is intra-linguistic, where, for certain languages, the default ϕ-featural value is precisely the one recruited for politeness. However, we must foreshadow the fact that the "bonus" second convergence is not necessarily mandated. While the pool of semantic defaults is universal, the question of which default(s) are selected from this pool as proxies of politeness is a highly language-specific matter. Languages may select a number proxy, a person proxy, or no selection at all.

With this disclaimer in place, let us consider typological data on the shapes of semantic defaults for the ϕ-categories of number/person (§3.3.1) and clusivity (§3.3.2).

3.3.1 Typological defaults for number and person

We will establish third persons (singular and plural) and impersonals as semantic defaults using the two diagnostics easily obtained from descriptive grammars, examining which ϕ-featural value surfaces in contexts of arbitrarity (§3.3.1.1–3.3.1.4) and/or as imposters (§3.3.1.5). These findings reinforce those above: third person singulars are person-neutral, and third person plurals and impersonals are number- and person-neutral.

3.3.1.1 3SG as default

In Ilokano (also known as Iloko, Northern Philippines, Malayo-Polynesian), the 3SG (ergative) clitic *=na* is the ϕ-featural value used to encode non-referential subjects. In the following prescriptions involving *rebbeng* 'should/must', 3SG *=na* appears even if the grammatical subject is second person (64a) or first person (64b). This is consistent with our English-based findings above in

(64) a. *Rebbeng=**na** laeng nga adda=ka=n iti denna=mi.*
should=**3SG**.ERG only LINKER be=2SG.ABS-EMPH OBL side=1EXCL.PL.ERG
'You should just be at our side.'

b. *Siak ti adda ka-rebbeng-an=**na** nga ag-bati.*
1SG ART be NOM-should-NOM=**3SG**.ERG LINKER 1SG-stay
'I am the one who should stay.' (Rubino 1997: 64-5)

For singular polite address, recall that 3PL *isuda* can be used in lieu of 2SG (along with 2PL *dakayo*, although using 3PL *isuda* conveys a higher level of formality) (Rubino 1997: 52).

Consider also the impersonal "passive" constructions in Iraqw (Tanzania, Cushitic). These are constructions in which the precise identity of the human subject is irrelevant. The language has a dedicated impersonal pronoun, *ta*, akin to English *one*, which may appear in subject position, as in (65a). The main verb is always conjugated for the third person singular masculine, as in (65a-c)—even if the subject is itself not third person singular, as in (65c).

(65) a. ***ta**-na hanúis tsat'i.*
IMPERS.COP-PAST give.**3SG.MASC**.PAST knives
'They gave knives.'

b. *tsat'i ki-na hanúis.*
knives **OBJ.IMPERS**.COP.PAST give.**3SG.MASC**.PAST
'Knives were given.'

c. *ti tlaahh nee fiisuuse.*
OBJ.IMPERS.1SG/2SG.F/1PL beat.**3SG.MASC**.PAST by thieves
'I/you(fem)/we are beaten by thieves.' (Mous 2007: 18)

The author further notes that "[i]f a person (pro)noun is expressed, it is the third person plural *ino'ín*", but no examples were given. No mention of politeness was found.

3.3.1.2 3PL as default

In Haya (Tanzania, Bantu), the 3PL subject affix *ba-* is used to expone the subject of impersonal constructions, which correspond to agentless passives in English (authors' own description). This

is illustrated in (66)[16].

(66) a. *ekitabo* ***bá****-ka-kí-gula.*
book **3PL**$_2$-REM.PAST-3SG$_7$.OBJ-buy
'The book was bought.'
b. *ebitooke* ***bá****-ka-bí-cumba.*
bananas **3PL**$_2$-REM.PAST-3PL$_8$.OBJ-cook
'The bananas were cooked.' (Duranti & Byarushego 1977: 49)

We see a bonus convergence between defaulthood and politeness in Haya. It is precisely this ϕ-featural value, 3PL, that is used in forming polite commands, directed towards a second person like (67). (Plain commands are marked with 2SG *o-*.)

(67) ***ba****-m-p'ómwóoño.*
3PL$_2$-1SG$_1$-give.salt
'Give me salt (polite).' (Lehman 1977: 147)

Givón (2015: 53-4) notes that a similar pattern of passive formation also pertains in another Bantu language, Kimbundu. Its passive is formed similarly with 3PL *a-*, even though the non-3PL agent may be included in an optional *by*-phrase (68).

(68) *nzua* ***a****-mu-mono (kwa meme).*
John **3PL**-3SG.OBJ-saw by 1SG
'John was seen (by me).'

We see that 3PL also surfaces in contexts of arbitrarity in an unrelated language, Paamese (Vanuatu, Malayo-Polynesian). It features a singular-dual-paucal-plural number system with a clusivity distinction. The author explicitly notes that "[t]he plural forms are also semantically unmarked", such that 3PL surfaces for unspecified subjects rather than the dual or paucal. (Even though this dissertation pretends that finer-grained multiplicities like dual or paucal do not exist, this may be evidence that out of all multiplicities, plural may be the most default one.) Sadly, we cannot establish a bonus convergence, as the author makes no mention of politeness.

(69) ***A****-lesi-e tăhosi.*
3PL.REAL-SEE-3SG 3SG.REAL.good
'It looks good.' (Crowley 1982: 80)

3.3.1.3 3SG, 3PL as defaults

Wandala (Cameroon/Nigeria, Chadic) features a singular-plural number system with a clusivity distinction, as (70) shows for the subject pronouns. (The diacritics mark tone, which may change depending on the phonological context the pronoun finds itself in.)

(70) Wandala subject pronouns (Frajzyngier 2012: 181)

[16]In the Haya examples, the subscript in the gloss line indicates noun class. *ba-* indicates the noun belongs in class 2 (plural humans), while *kí* and *bí* indicate the noun belongs in classes 7 and 8 (respectively, singular and plural artifacts and things) (Duranti & Byarushego 1977: 9).

	SG	PL
1EXCL	*yè/yà*	*mà/myà*
1INCL	–	*ŋà*
2	*kà*	*kwà*
3	*à*	*tà*

Of these, the 2SG, 3SG, 3PL, and 1EXCL.PL are explicitly described to be used for arbitrary reference (Frajzyngier 2012: 521-2). Such uses are illustrated below, in context whenever possible, to ensure that arbitrary reference was in fact intended by the speaker. We will first start with the arbitrary uses of the third person forms.

(71) is taken from a conversation where two people in the older generation are expressing their grievance about modern dating practices, where the speakers' opinion is that boys court girls inappropriately (Frajzyngier 2012: 522/680-2 ex. 42, 51-2). Arbitrary use of 3SG *á* can be seen in (71b), where it refers to an unspecified person. The corresponding 'one' in the translation line is the author's, and the author also explicitly states that its "referent may include the speaker and hearer" (p. 522). This makes sense: any elder would be in a position to admonish such boys, including either elder speaker or hearer.

(71) a. *bà gdzàr-á-mì zál-àhà nà kínì wá tá gdzàr-á*
FOC child-GEN-1PL male-PL DEM FOC COMMENT 3PL bother-3PL-TARGET-GOAL
ìrè má tsà-rə̀-tsé gə̀ gdzàr-á gyálà-hà...
head HYP rise-3PL-rise to child-GEN female-PL
'it is the boys, who bother the girls...'

b. *ítàr ŋánnà má ské* ***á*** *má bà-trá wè gə̀ gdzàrà zál-àhà ŋánnà?*
3PL DEF HYP like.that **3SG** HYP say-3PL what to small male-PL DEF
'If they are like that, so what does one tell those boys?'

c. *tá kàtà bà ŋgùɗír kwá tá kàtà bà tá bàdzà əgdzàrà-hà bí.*
3PL want FOC promiscuity and 3PL want FOC 3PL spoil child-PL completely
'They want promiscuity and spoil the girls completely.'

The third person plural is used for expressing arbitrarity as well. (72) is taken from an introduction to a story about three people that did not originate with the speaker, the *Dead father's story* (Frajzyngier 2012: 521/658 ex. 2-3). The originator of the story is unknown, and is marked via 3PL *tà* in (72b).

(72) a. *yè s-á gdzà nàrwàndzà án w-á wándàlà.*
1SG come-GOAL small story ASSOC mouth-GEN Wandala
'I brought a small story [of] Wandala [from someone].'

b. *yò* ***tà*** *s-á lv-á mdə̀ kíɠyé.*
well **3PL** come-GOAL story-GEN people three
'They [this someone] brought a story of three people.'

(73) shows another arbitrary use of the third person plural. It is taken from a story in which a man boasts about his counting abilities (Frajzyngier 2012: 636 ex. 23-24b). In (73b), the unspecified subject 'one' (translation taken directly from the author) is marked with 3PL *tà*. (Note that, in (73c), 'nobody' is anteceded by 3SG *á*, further illustrating the defaulthood of the third person singular.)

(73) a. *má bání lìsáfì-á dùksə̀ bàɗə̀mmè*
HYP concern calculation-GEN thing all
'If it concerns any kind of calculation,'

b. *ɓákà wrà* ***tà*** *tàttàyà à j-ú g-íyà ɓákà.*
never person **3PL** search 3SG surpass-VENITIVE to-1SG never
'One does not look for a person to surpass me.'

c. *má bà kəɮà kùlà wá* ***á*** *jì-y-ú k gə̀ ìyà.*
HYP FOC count calculus COMMENT **3SG** surpass-1SG-VENITIVE NEG to 1SG
'If it concerns counting, nobody surpasses me.'

Wandala also features a second person impersonal pronoun, akin to English *you*. It is the second person singular *ká*. (74) was given without context (Frajzyngier 2012: 345 ex. 48), but the corresponding translation of *ká* as 'one' is taken directly from the author.

(74) ***kà***-*ʃkwà-hè cùkwá ŋgúɗì lóktù* ***ká*** *sà-wà.*
2SG-last-PUNCTUAL a little time **2SG** come-VENITIVE
'One has to wait a little bit before coming.'

(75) was taken from an interview-like conversation about contemporary marriage (Frajzyngier 2012: 676 + 522/678: ex. 26, 27, 31). The speakers are expressing grievance about the trend of girls staying unmarried, with their mothers supporting this decision, therefore undermining the wishes of their fathers. (75c) expresses the speaker's powerlessness about the state of affairs, with 2SG *ká* being used as an impersonal. The author explicitly states that 2SG *ká* "may include both the speaker and the listener" (p. 522), which makes sense as both participants of this conversation are commiserating with one another.

(75) a. *dàdà á kàtà gdz-á-rà á dəm mbá žilé.*
father 3SG want daughter-GEN-3SG 3SG go house man
'A father wants his daughter to marry.'

b. *è màmá séy á hàyá bá-úná à wwày-nà rvòd-á-r-nà.*
eh mother then 3SG love FOC-DEM 3SG love-3SG heart-GEN-3SG-DEM
'and the mother wants what her [the girl's] heart wants.'

c. *áyò ə́gdzar-à kínnà wá má* ***ká*** *bà-trá lvà à bà-itáré á*
now child-PL FOC COMMENT HYP **2SG** talk-3PL speech 3SG say-3PL 3SG
s-á-nà zàmàné.
come-3SG time
'And now, children, when you talk to them, they say it is modern times that brought it.'

Judging from the singular morphology, it is likely that the route to impersonality taken by 2SG *ká* is one of distributive quantification, not plural predication.

The last ϕ-featural value that may be used arbitrarily is 1EXCL.PL. It is the only counterexample to the generalization which will be established in §4.2.2 that I know of, where inclusives are default forms, not exclusives. However, try as I might, I could only find one example of such a use, presented in (76b). The context of (76) is the same as that of (71) above, where two of the older generation are both expressing their grievances about modern dating (Frajzyngier 2012: 521/691: ex. 109-110).

(76) a. *tà mùsəlmá tsà t-á-m səlàwré!*
3PL Moslems yet 3PL-PRED-in pants
'They are Moslems, and yet they wear pants!'

b. *də́-ŋgré səlàwré gdz-à-rà mùksè kínì səlàwré...*
father-**1EXCL.POSS** pants child-GEN-3SG female FOC pants
'One's father (wears) pants, even his daughter (wears) pants.'

The use of the exclusive here cannot necessarily be attributed to the referents being in a different ethnic group than the speaker, as many Wandala are Muslims (Frajzyngier 2012: 4). However, the author notes that for (76b), the "actual speaker may in fact not be the part of the group involved" (p. 521) which is plausible given the context of the utterance: (76) is taken from the same story as (71) where two elders are expressing their grievance about modern dating practices. Thus, the speaker's attitude is one of disbelief, disdain, and disassociation, and it may be reasonable to conclude that the exclusive was used conventionally in this instance, and not as a default. Judging from its rarity and the unclear context, I will not consider the exclusive plural as a ϕ-featural value that appears in contexts of arbitrarity.

Thus, Wandala features a panoply of ϕ-featural values that appear in contexts of arbitrarity: 2SG, 3SG, and 3PL. Despite this panoply, none of them are identical to the proxy of politeness. (But this is fine.) In Wandala, it is 2PL which is used for singular polite address, as is illustrated in (77), which is addressed to a person older than the speaker.

(77) *yò ʃóyì ŋánnà wàtsè* ***kwà*** *d cìn-ná kúr nà nó-ŋwàné lvà wàndàl...*
well chat DEF FUT **2PL** fut hear-3SG 2PL DEM PRES-DEF language Wandala
"Here is the chat that you (POLITE) will hear in Wandala...' (Frajzyngier 2012: 539 ex. 105)

3.3.1.4 Dedicated impersonal forms as defaults

Here, we examine a handful of languages with dedicated impersonal forms, and show that these exhibit properties of semantic defaults *and* are recruited for polite address.

Recall from (21b) that the impersonal form *on* can be used for polite address in French, but only in extremely formal contexts. In colloquial French, impersonal *on* is number- and person-neutral, being the pronoun with the widest range of coverage over different ϕ-featural combinations. *On* can be used to refer to a group of mixed grammatical persons: first and third in (78a); first, second and third in (78b).

(78) a. *Marie et moi,* ***on*** *est gentille-s.*
Marie and me IMPERS COP.3SG nice.FEM-PL
'Marie and I, we are nice.'
b. *Marie et toi et moi,* ***on*** *est gentils.*
Marie and you and me IMPERS COP.3SG nice.MASC-PL
'Marie and you and I, we are nice.' (Hénot-Mortier, p.c.)

French *on* may also be used to imposter for first person plural (Laberge & Sankoff 1979). A specific example is provided by Harbour (2016: 245-6): (79) is a passage from Proust (Harbour's translation) which describes a lost cake; in it, *on* is substituted for *nous* without change in reference.

(79) *Au moment où* ***nous*** *arrivions à la maison, maman s'aperçut qu'****on*** *avait oublié le saint-honoré...*
the moment that we reaching to the house, Mother noticed that-IMPERS had forgotten the Saint-Honoré
'Just as we were reaching the house, Mother noticed that **one (=we)** had forgotten the gâteau Saint-Honoré...'

In Yiddish, the impersonal pronoun *men* is used to expone arbitrary or generic subjects, as the naturalistic examples show in (80a-b).

(80) a. *Vos zhe tut* ***men****, az* ***me*** *tor nisht esn tsu fil fets?*
what then do IMPERS if IMPERS ought not eat too much fat
'What does one do, if one ought not eat too much fat?'

b. *vegn dem shraybt* ***men*** *shoyn in der britisher medyle.*
about this write one already in the british media
'People are already writing about this in the British media.' http://web-corpora.net/YNC/

me(n) is indeed number- and person-neutral, as (81) shows.

(81) ***Me*** *hot dos shoyn gepruvt.*
IMPERS has that already tried
'Someone (me, you, s/he, we, they) has already tried that.' (Adam Albright, p.c.)

Recall from (22) that *me(n)* can be used in polite contexts.

To end our tour of languages with dedicated impersonals as semantic defaults, I will mention Kelabit (Malayo-Polynesian, Borneo). In Kelabit, *narih* is the form used in contexts of arbitrarity, exponing the subject of generic and prescriptive statements. (The indication of 'generic' in the translation line is taken directly from the author.)

(82) a. *Buken* ***narih*** *tudo ngi rumaq...*
NEG IMPERS sit at house
'We (generic) didn't just sit around at home...'

b. *Napuh neh* ***narih*** *lem rumaq* ***narih****.*
sweep PART IMPERS in house IMPERS
'We (generic) swept in our houses.'

c. *Mey meman berek neh* ***narih****.*
go AV.feed pig PART IMPERS
'We (generic) went to feed the pigs.'

d. *Getebp-en tukung* ***narih****, rengaq* ***narih*** *na'am liang kelaboq ih.*
bite-UV.IRR mosquito IMPERS if IMPERS NEG under net PART
'You (generic) will get bitten by mosquitoes if you don't sleep under a net.' (Hemmings 2016: 181-2)

narih also has the widest range of reference over distinct grammatical persons: it can be used to refer to the addressee(s) (83a-b), the addressee and others (83c), or the speaker (83d). It is therefore number- and person-neutral.

(83) Kelabit *narih* covers all grammatical persons

a. *Belajar tu'uh-tu'uh* ***narih*** *keyh... kedeh.*
study REDUP-true IMPERS PART say.3PL
'Make sure you (the addressee) study hard, ok, they said.'

b. *Kuman dooq-dooq* ***narih!***
AV.eat REDUP-good IMPERS
'Bon appétit (to a single addressee or group)!'

c. *Ngudeh* ***narih*** *na'am mey ngalap kayuh toq?*
why IMPERS NEG go AV.fetch wood first
'Why didn't you (addressee+contemporaries) go and get wood first?'

d. *Naruq nuk kereb tu'-en* ***narih****.*
AV.do RELATIVE can do-UV.IRR IMPERS
'Do what I can do.' (Hemmings 2016: 181-2)

Again, we see two diagnostics of defaulthood, impostering and arbitrariness, converging on fourth person *narih* in Kelabit. There is a further convergence with politeness: recall from (20) that in polite imperatives, it is precisely *narih* which expones the grammatical subject.

To summarize, when we additionally consider languages with dedicated impersonal pronouns, these also emerge as a semantic default. Furthermore, impersonals are default in the extreme: it is not only person-neutral, but also number-neutral.

3.3.1.5 Third persons as imposters

In several European languages, we find that third-person expressions have grammaticalized into first or second person pronouns (e.g. Romanian variations on *dumneavoastră*, now respectful second person, from the third person DP *Dumnia Voastra* 'your.PL grace', see Wang 2023: §4.3). Currently, in Brazilian Portuguese, the third person expression *a gente* 'the people' is replacing the 1PL pronoun *nos* (Zilles 2005; Pacheco et al 2024). Conversely, we do not find diachronic evidence for grammaticalization in the other direction, where first or second person pronouns evolve into third-person expressions.

Kinship terms can be used pronominally cross-linguistically. I briefly illustrate with Fehan Tetun, which allows third person imposters to take on both second and first person reference. Imposters consist of kinship terms like *ina* 'mother', metaphorically extended kinship terms like *am-bei* 'ancestor', the common noun *na'i* 'noble', highly common proper names such as *Manek* and *Bete* (regardless of the referent's actual personal name). These can all be used in lieu of a second person pronoun. Furthermore, instead of using a first person pronoun, some speakers refer to themselves with the title by which they prefer to be addressed with. (84) is adapted from van Klinken (1999: 140), showing this with the third-person term of address *ina* 'mother'.

(84) ***Ina*** *dakar oan hawa'i hanesan oan té.*
mother look.after child adopt resemble child give.birth
'Mother (=I) looks after adopted children like her children by birth.'

This impostering of kinship terms for first singular also pertains in caretaker speech in English, *e.g.* a parent promising her bullied child, *Mommy will see that you're avenged* (Wills 1977), and maybe more widely (Blackwell 2007).

3.3.2 Typological defaults for clusivity

Readers may recall that inclusives figured prominently as proxies of politeness in §2.1.3. Since English does not make a morphological clusivity distinction, we examine languages which do in order to diagnose defaulthood within the inclusive–exclusive opposition. In such languages, there is a distinction between two types of first person plural. The *inclusive first person plural* pronoun refers to the group comprising the speaker, addressee, and any number of others (including zero). In contrast, the *exclusive plural* denotes the group comprising the speaker and any number of others (including zero)—but crucially excluding the addressee.

This section will establish that it is the inclusive, not the exclusive, which passes diagnostics for semantic defaulthood. To my knowledge, this constitutes a novel observation in the ϕ-featural literature[17]. §3.3.2.1 shows that it is the inclusive which appears in contexts of arbitrarity. §3.3.2.2 shows that imposters may take on the morphological shape of inclusives, but not exclusives. Lastly, §3.3.2.3 traces a diachronic affinity between inclusive and impersonal

[17]The only potential counterexample was flagged above in (76), where the exclusive plural was described as one of the ϕ-featural means for impersonality in Wandala.

pronouns in Indo-Aryan.

3.3.2.1 When does the inclusive not quack like an inclusive? When it quacks like an impersonal

Here, we examine languages without a dedicated fourth person pronoun, and ask, what form is used to expone the subject in contexts of arbitrarity (e.g. generic, prescriptive, public utterances)? Below, we see that it is the inclusive that approximates the fourth person.

In Tukang Besi, it is the inclusive pronoun which is used to realize the subjects of generic statements. In (85), the reduced realis subject allomorph of 1INCL *ikita, to-*, is used.

(85) a. ***To**-[m]o-ha'a-ke?*
1INCL.PL.R-VERBALIZER-why-3OBJ
'What shall we do about it?'
b. *Ara **to**-sai te humbu...*
if **1INCL.PL**.R-make ART k.o.basket
'If you make a basket,...'

Ainu does have a fourth person form *-a(n)*, but it is not a dedicated form: *-a(n)* also expones the inclusive subject marker, which is realized as *a-* in transitives (86a) and *-an* in intransitives (86b).

(86) a. ***A**-e-kore.*
1INCL-2SG-give
'We give you something.'
b. *Itak-**an**.*
speak-**1INCL**
'We speak.' (Shibatani 1990: 56)

Given that *-a(n)* is also recruited for politeness (cf. (17b)), a clause with *-a(n)* in subject position is actually ambiguous between an inclusive and a polite reading, as in (87).

(87) ***A**-en-kore.*
1INCL/IMPERS-1SG-give
'You (POLITE) give me something.'
(*Literally.:* '<u>We</u> give me something.') (Shibatani 1990: 56)

Naturally, *-a(n)* is used for generic (88a) and prescriptive statements (88b). These are non-referential subjects.

(88) a. *Tepeka paye-**an** yak Sat ta paye-**an**.*
here go-IMPERS if Saru to go-IMPERS
'If one goes here, one goes to Saru.'
b. *Tan chep anakne **a**-satke wa **a**-eiwanke-p un.*
this fish TOPIC IMPERS-dry and IMPERS-use-thing is
'This fish is a thing that we dry and use.' (Shibatani 1990: 56)

Hence, Ainu is one language where one morpheme is shared three ways between non-referential, inclusive, and polite uses[18].

[18]The Ainu utility of *-a(n)* does not stop at the impersonal-inclusive-honorific trinity: it also has verbal uses.

Fehan Tetun is another such language. Recall the polite inclusive *ita* from (8), repeated below as (89). Its inclusive use is illustrated in (90).

(89) *Ama Liu, ha'u k-mai k-amán* ***ita****...*
father Liu 1SG 1SG-come 1SG-accompany **1INCL.PL**
'Father Liu, I have come to accompany you (POLITE)... '
(*Literally.:* 'I have come to accompany both of us...')

(90) ***Ita*** *na'in hity ne'e,* ***it*** *ha-la'e hotu ó la'en ne'e.*
1INCL.PL CLS seven this **1INCL.PL** make-husband all 2SG husband this
'We seven (including you), we will all marry your husband.' (van Klinken 1999: 112)

ita is also used in normative statements such as (91), as "explanations of how things are normally done, or used to be done, or when describing hypothetical situations" (van Klinken 1999: 112).

(91) ***Ita*** *feto la'o ha'i.* ***Ita*** *iha uma karian.*
1INCL.PL woman walk not **1INCL.PL** LOC house work
'We (generic) women don't go out. We (generic) work in the house.' (said to a male)
(van Klinken 1999: 112-3)

A related language, Banggai (Indonesia, Malayo-Polynesian), also features the same trinity of default-inclusive-polite. The inclusive plural *ikita* is used for polite address, and for impersonal reference (van den Bergh 1953: 44-5).

In Sino-Tibetan, the same trinity of default-inclusive-polite can also be found. For Limbu (Bhutan/India/Nepal). van Driem (1987: 220) shows that *a-* is the only verbal agreement morpheme that is ambiguous between non-referential and referential uses. Its non-referential uses are illustrated in (92).

(92) Limbu impersonal inclusive (van Driem 1987: 218-9)

a. *kh?o:* ***a****-ya:-i:* ***a****-n-ya:--i:?*
there **1INCL**-scare-Q **1INCL**-NEG-scare-NEG-Q
'Does it get scary there or doesn't it?' (lit. 'Will it scare us there or won't it?')

b. *hkke: co:k-m-lle* ***a****-si!*
like.that do-INF-SUBORD **1INCL**-die
'You'll die if you do it that way!'

c. *canʉkwa ca-m-lle k:b-ba* ***a****-bo:.*
canʉkwa eat-INF-SUBORD mute-NOMLZ **1INCL**-become
'If you eat *canʉkwa*, you'll become mute.'

d. ***a****-ya!*
1INCL-strike
'It [a projectile] will strike us!'

Recall from (40), repeated as (93) below, that the language features honorific inclusivity.

(93) Limbu (adapted from van Driem 1987: 221)

a. *hn* ***a****-dzo:k-si-rʉ* ***a****-ya:k-si?*
what **1INCL**-do-DU-PRES.GER **1INCL**-be-DU
'What are you (POLITE) doing?'
Literally: 'What are we two doing?'

Notably, these are not just *any* verbal uses, but precisely those which are non-referential: the spontaneous and the passive. I cover these additional uses of *-a(n)* later.

b. *hn* ***a****-dzo:k-Ø-l* ***a****-ya:k-Ø?*
what **1INCL**-do-PL-PRES.GER **1INCL**-be-PL
'What are you all (POLITE) doing?'
Literally: 'What are <u>we all</u> doing?'

Dramatically, in languages without dedicated forms for either fourth person or inclusive, we see that approximated inclusives are the forms used for utterances without specific interlocutors. I detail this for Caddo (with its composed inclusive) and for Tongan (with its pseudo-inclusive).

Caddo

Caddo does not have a fourth person pronoun or a clusivity distinction. Its pronoun system is comprised of pronominal prefixes, which make distinctions between first, second, third person, across the numbers singular and plural. (Pronouns are also distinguished for realis and irrealis, but I only present the realis series here.)

(94) Caddo realis pronominal prefixes (Melnar 1998: 48)

	AGENT	PATIENT	DATIVE
1	*ci-*	*ku-*	*ku-*
2	*yah'-*	*si-*	*si-*
3	*Ø-*	*Ø-*	*(n)u-*
DEFOCUSED	*(y)i-*	*ya-*	*yu-*

An additional grammatical person, "defocused third person", is described. It is termed as such because it is the pronoun used for referencing non-protagonists in narratives. In this story, Turtle is the protagonist. He and several other creatures are looking for water in a drought. When Turtle gets stuck in a log, the other creatures trample on him.

(95) a. *kú si dí ha 'imay hús sa- ya'k- 'ini- 'a'*
just time this ADJ big MIRATIVE IRR wood lying be
'After a while, there was a big log lying!'
b. *nátti' yika=hihy -ah*
there get.stuck PERF
'There he (Turtle) got stuck...'
c. *natti'* ***yi-*** *'ini- bi'=sak -ah*
then **DEFOC** horizontal.surface step.on.forcefully PERF
'Then they (the other creatures) trampled on him (Turtle)...'

In (95b), the lack of overt pronominal morphology indicates that the referent is a 3rd person, where Turtle is the focus of the narrative. In (95c), the tramplers of the stuck Turtle are referenced by the defocusing *yi-*. This way, the speaker maintains focus on Turtle, the protaganist, throughout the narrative.

I will argue that what has been described as "defocused third person" is actually a fourth person. To start, these forms can only mark human referents (Melnar 1998: 50), a requirement also carried by impersonal pronouns. (The creatures above are anthropomorphized for the narrative.) Furthermore, these forms have three further impersonal-like functions in Caddo: i) naming human-oriented common nouns, ii) marking an obligatory number-neutral distributive, iii) politeness, and iv) composing the inclusive.

Starting from i), we see that *yi-* ∼ *ya-* appears in names denoting common nouns which are literally constructed as entities with a typical component that defines the particular instrument, location, patient, agent, or action. This prototypical component is translated into English

frequently via the impersonal *one*. Given the humanness requirement of *yi-* ∼ *ya-*, it is expected that (96) are all common nouns involving human agents.

(96) a. *nak* ***yi-*** *ki=sáw' -'unah (=nakihsáw'unah)*
INSTR.SUBORD **DEFOC** scrape MIDDLE
'razor (lit. that with which one scrapes oneself)'

b. *bak'a'nuh kúk* ***yi-*** *yán- yási' -sa (=bak'a'nuh kúkíyánná-sa')*
word LOC.SUBORD **DEFOC** CAUS roam IMPF
'courthouse (lit. where one causes words to roam)'

c. *nit* ***yi-*** *kátta- bínik (=nídikáttabínih)*
PAST.GENERIC.PARTICIPLE **DEFOC** liquid move.circularly
'fried corn mush (lit. liquid that one has moved circularly, stirred)'

d. *nayt* ***ya-*** *ʔi=yan -hah (=naytdayánhah)*
GENERIC.PARTICIPLE **DEFOC** catch HAB
'policeman (lit. one who catches people)'

e. *kak* ***yi-*** *ka=dís (=kakikadis)*
SUBORD **DEFOC** wash
'(for one) to wash' (Melnar 1998 citing Chafe 1990: 61-64)

Use ii) is the defocused person's role in marking number-neutral distributivity. It is not, to my knowledge, a role of defocused person that has been pointed out in the literature on Caddo before. Distributive affixes on the verb serve to express multiplicity of events, and Caddo distinguishes *individuative vs. varietive* distributivity: while the individuative *haka-* expresses that different entities are involved in the events, the varietive *kaki-* expresses that the entities are varied in kind. Varietive distributivity, expressed with the help of a defocusing form in (97a), is incompatible with singularity and plurality; this recalls another property of fourth person, its number-neutrality. This is in contrast to the individuative, which appears with plural marking *wa-* in (97b).

(97) a. *kak* ***yu-*** ***kaki-*** *t- 'i- ya'ah (=kah'ukakít'í'ah)*
SUBORD **DEFOC.DAT VAR** APPL dance be
'all kinds of dances (lit. be different kinds of dances)' varietive

b. *nít* ***haka-*** *wa- 'án -ah (=nítakáw'ánah)*
PAST.GENERIC.PARTICIPLE **INDV** PL choose PERF
'he chose them, caught them (e.g. fish)' individuative

Use iii), marking politeness towards in-laws, was already covered in Chapter 2. This is repeated below in (98), where it is used for both polite address (98a) and polite reference (98b).

(98) a. *dikat-* ***yi-*** *'adih- hah? (=dikadiidihah)*
what- **IMPERS**- go- HAB
'Where are you (POLITE) going?'

b. *dikat-* ***yi-*** *'adih- hah? (=dikadiidihah)*
what- **IMPERS**- go- HAB
'Where is he/she (POLITE) going?'

Lastly, we arrive at iv), the role of defocused person in composing the inclusive. Since Caddo does not make a morphological distinction in clusivity, the language expresses it from the combination of other existing elements[19]. However, the elements selected to approximate

[19]For more instances of composed inclusives, see Harbour (2016: §5.2.1). Another phenomenon of a composed category is the dual; see Harbour (2020) on "Frankenduals", the phenomenon where dual meaning emerges from

the inclusive are not random: the inclusive is constructed with none other than the defocusing suffix. When *yi-* is combined with the animate dual *wiht-* or the animate plural *wa-*, the result is inclusive dual or inclusive plural, respectively, shown in (99a) and (99b). (Below, the underlying *yi-* in the verb is not the defocusing prefix, but part of the verb *yi=bahw-* 'to see', formed with a preverb *yi=* (Chafe 1976: 45-6). I deviate from the author by glossing the stem and its preverb together as one morpheme.)

(99) Caddo inclusive: composed from inclusive and number (Chafe 1990: 67, adapted)

a. ***yi-*** ***wiht-*** *yibahw- nah.* *(=díwísbáwnah)*
DEFOC- **DU-** see- PERF
'We (me and you) have seen it.'

b. ***yi-*** ***wa-*** *yibahw- nah.* *(=díiwabáwnah)*
DEFOC- **PL-** see- PERF
'We (me and you and others) have seen it.'

To substitute impersonal *yi-* with the general first person marker *ci-* is to produce an exclusive meaning, as (100) shows.

(100) Caddo exclusive: composed from general first person and number (Chafe 1990: 66, adapted)

a. ***ci-*** ***wiht-*** *yibahw- nah.* *(=cíwísbáwnah)*
1- **DU-** see- PERF
'We (me and another, but not you) have seen it.'

b. ***ci-*** ***wa-*** *yibahw- nah.* *(=cíiwabáwnah)*
1- **PL-** see- PERF
'We (me and others, but not you) have seen it.'

What kind of core meaning might tie together these disparate uses of the defocusing prefix covered in i)–iv)? Uses i) and ii) show that the defocused person carries a humanness requirement and is number-neutral, respectively—two hallmarks of an impersonal pronoun. Uses iii) and iv) show that it is used to approximate politeness and inclusivity, respectively—two proxands of the impersonal that are supported cross-linguistically. Hence, the defocused person is not a subtype of third person, as its moniker suggests, but rather an impersonal or fourth person.

Tongan

Tongan (Tonga Islands, Oceanic) features a "pseudoinclusive". Since Tongan is not a minimal-augmented language, there should be an obligatory gap for the "inclusive singular", a dedicated form for the speaker-hearer dyad. However, this gap is filled by the form *te* (preposed form, its postposed form is *kita*).

(101) Tongan preposed pronouns

	SG	DU	PL
1EXCL	*ou*	*ma*	*mau*
1INCL	***te***	*ta*	*tau*
2	*ke*	*mo*	*mou*
3	*ne*	*na*	*nau*

However, it is precisely this form (given the non-standard gloss 1INCL.SG) which is used for the subjects of generic statements (102), suggesting its defaulthood. Furthermore, it is the form

particular compositions of singular with plural.

used for social deixis (103).

(102) *'Oku 'ikai totonu ke* ***te*** *tokanga pē kiate* ***kita.***
pres neg right that 1INCL.SG attend only to 1INCL.SG
'It is not right that one should attend only to oneself.' (Harbour 2016: 246, citing Churchward 1953)

(103) *'Okú* ***te*** *mā 'aupito* ***kita*** *'i he'enau angafaí.*
pres 1INCL.SG ashamed very 1INCL.SG loc their conduct
'I (HUMBLE) for my part am very ashamed of their conduct.' (Harbour 2016: 247, citing Churchward 1953)

In Bajau (Indonesia, Malayo-Polynesian) the pronoun *kita* is already doing double duty as the inclusive and (predictably) as the singular second person honorific. For "re-inclusivization", it is the impersonal form *aha* 'man, person' that is being used, to emphasize inclusive meaning where *kita* would have been ambiguous between inclusive and polite address.

(104) *Sabbar né* ***aha,*** *emma. Poré ja* ***kita*** *lagi saloh.*
be.patient EMPHATIC **person** mother go only **1INCL.PL** again tomorrow
'Let us have patience, mother. (You) Just go once again tomorrow!' (Verheijen 1986: 19; gloss from Lichtenberk 2003: 272 except for *sabbar; né* from Verheijen 1986: 107; 20)

The impersonal use of *aha* may be a change-in-progress. It does have referential uses such as *aha toa* 'old people', *dakau aha* 'a man', *aha kaséasé* 'poor people' (Verheijen 1986: 16, 20, 36). However, non-referential uses of *aha* are also attested by the following examples, where it expones unspecified subjects.

(105) *Nggai lo-lōmba ké* ***aha*** *iru.*
not LO?-compete.running Q **person** that
'It was impossible for people not to run.' (Verheijen 1986: 38, gloss my own)

(106) *Dadi nia* ***aha*** *pabara bahwa nia dakau saudagar mandoré.*
and to.be **person** announce C to.be a merchant yonder
'Actually people said there was a certain merchant.' (Verheijen 1986: 42, gloss my own)

Another Oceanic language, Tikopia (Solomon Islands) also innovated a dedicated impersonal form like English 'one', *kita*. It comes directly from the Proto-Austronesian inclusive form **(k)ita* (Firth 1985: 188). The author does not mention polite pro-forms.

Notably, many of the languages discussed above each have a place in our typology of polite pronouns. The inclusive is already the proxy for politeness in Fehan Tetun (8); Ainu (17b); Tukang Besi (50); Limbu (40); Caddo (98a); Tongan (103); and Bajau (104).

Further examples of the impersonal–inclusive connection outside our typology of polite pronouns can be given. Storch (2014: 151) gives the following minimal pair for Luwo (Sudan, Nilotic). (107a) with the exclusive plural expresses a particular group's habit of eating guava, while (107b) with the inclusive plural is a generalized or impersonal construction. Generally, the Luwo inclusive "expresses common practice and general truth" (Storch 2014: 154).

(107) a. *jòwááfà bɛ́ɛ́dh=ɛ́ cám* ***wá****=ù-cʌ́mɔ́.*
guava be=3SG eat **1EXCL.PL**=IMPF-EAT.AP

'We normally eat guava.'

b. *jòwááfà béédh=é cám* ***òn****=ù-cámɔ́.*
guava be=3SG eat **1INCL.PL**=IMPF-EAT.AP
'One normally eats guava.'

The author notes the synchronic similarity between the inclusive and the third person, and proposes a diachronic pathway between the two categories, where third person is lexicalized into an impersonal. This process is also documented for other languages in Eastern Nilotic (Dimmendaal 1991).

In Acehnese (Indonesia, Malayo-Chamic), the inclusive pronoun *geutanyoe* (clitic form *ta*)[20] is used in contexts of arbitrarity. It "is often used as an indefinite pronoun which could refer to anyone, the speaker or the addressee included. This is particularly for stating general truths about human activity or behavior" (Durie 1985: 120). (108a) describes a general practice. (108b) was said from one man to another man, and the conventional meaning of the inclusive is obviously not intended here.

(108) a. *bruek* ***ta****=peugöt keu=aweuek.*
coconut.shell **1INCL**=make to=ladle
'Coconut shells are made into ladles.'

b. *meunyo ka=lheueh* ***ta****=meukawên ureueng=inöng=****geutanyoe****=nyan*
if 2.INTIMATE=finish **1INCL**=marry person=woman=**1INCL**=that
*éh sa=jan=****geutanyoe****.*
sleep one=time=**1INCL**
'Once you are married your wife sleeps with you.' (Durie 1985: 120, ex. 5-39; 5-40)

An example of 'parental *we*' is also given, where the intended referent of the inclusive is the addressee, as in (109).

(109) ***geutanyoe*** *ka=chik, bek* ***ta****=peuranguy lagee=aneuk=manyak.*
1INCL 2.INTIMATE=adult don't **1INCL**=act way=child=infant
'We're grown up now; let's not act like a child.' (Durie 1985: 120, ex. 5-42)

Above, we have established a proxy-proxand relationship of another kind, that has nothing to do with politeness or social deixis: the inclusive approximating the impersonal in languages without a dedicated impersonal pronoun. Why is this proxy-proxand relationship possible? I will later argue in §4.2.2 that this is because the two pronouns have highly similar meaning by dint of sharing a plural core.

3.3.2.2 Inclusives as imposters

In English, inclusive *we* is used stylistically to refer to the addressee. This is the so-called 'nurses' *we*', as in (110a), which can also be found in non-nurse contexts, such as a parent patiently instructing their child in (110b). Similarly, there also exists the 'intimidation *we*', where a gangster might utter (110c) shortly before, or during, the drawing of arms. (110d) features the use of the inclusive for expressing affection or solidarity, where the human owner of a pampered cat is using the inclusive even though a human owner cannot also have fur (let

[20]For anyone who was expecting the Acehnese inclusive plural to resemble the Proto-Austronesian inclusive **(k)ita*, *geutanyoe* is formed from *geuta* (cognate to *kita*) and the proximal demonstrative *nyoe* (Durie 1985: 131).

alone fur shared with a cat).

(110) a. (nurse to patient:) **We** (=you) need to be patient.
b. (parent to child:) **We** (=you) need to eat **our** (=your) breakfast now.
c. (gangster to civilian:) Do **we** have a problem here? (Do you have a problem that you would otherwise want to take up with me now?)
d. (cat owner to beloved cat:) Those nasty neighborhood cats won't ruffle **our** (=your) fur again.

While all the contexts in (110) feature increased social distance between speaker and interlocutor, they do not indicate the type of increased social distance typical of politeness. Rather, the increased social distance seems to signal antagonism or impatience as in (110a-c) or affection as in (110d), which are *not* characteristic of politeness. However, inclusive *we* also has "true" polite uses in service contexts, such as the 'waiters' *we*' as in (111).

(111) (Waiter to diners:) Hi, my name is Persephone.
I'll be taking care of **us** (=you) tonight... What are **we** (=you) thinking for drinks?

Impostering in the reverse direction is not possible: the second person pronoun cannot be used to refer to the speaker-hearer dyad. In (112a), an Olympics coach is telling her athlete that both of them have to be patient in waiting for the final results; *you* cannot be used in place of *we*. In (112b), on the day of a parent-and-child egg-and-spoon relay race, the parent is expressing to their child that both of them must eat both their breakfasts for joint optimal performance; *you* cannot be used in place of *we* either. In (112c), the gangster cannot use the second person pronoun for intimidation; rather, (112c) sounds like an offer of help.

(112) a. (coach to athlete:) #**You** (=we) need to be patient.
b. (parent to child:) #**You** (=we) need to eat **your** (=our) breakfast now.
c. (gangster to civilian:) #Do **you** (=we) have a problem here?

In several Mixtec languages, the inclusive is frequently used in soliloquy for stylistic purposes. This is the case for Coatzospan Mixtec (Small 1990), where the 1INCL.PL forms are *ō/kō* (subject/object). In (113a), the single owner of a cornfield is very angry to discover that her cornfield has been eaten. We can be sure that the reference is singular, because the owner is cross-referenced repeatedly by the 3SG feminine pronoun *tún* in the later utterance (113b).

(113) a. *shoó ndu kaehshí i itú* ***kō****?*
who Q eat 3.UNSPECIFIED cornfield **1INCL.PL**
' "Who has been eating our cornfield?" '
b. *sá kahan tún da násheé tun nü vīhi tún kān.*
so speak 3SG.FEM when arrive.back 3SG.FEM face house 3SG there
'she said when she arrived back at her house.' (Small 1990: 448)

The same phenomenon of using the inclusive to refer to oneself is found in Alacatlatzala Mixtec (Zylstra 1991: 119) (114) and Diuxi-Tilantongo Mixtec (Kuiper & Oram 1991: 341) (115).

(114) a. *tā ndīvahyí shíhī ní rí sokō.*
and coyote die INTENSIFIER it hunger
'But the coyote was very hungry.'
b. *míkía kwahan ní ñānī* ***yó*** *yóhō ta shíhī ní*
where go INTENSIFIER brother **1INCL.PL** thsi and die INTENSIFIER

yó *sokō?*
1INCL.PL hunger
‘ “Where has this brother of ours gone for so long, and we are dying of hunger?” ’
(Zylstra 1991: 169)

(115) a. *kúvete konexó.*
be.happy rabbit
‘The rabbit was happy.’
b. *n-kakú-ŕ! ná xín ndukú* ***ró*** *shá káá* ***ró.***
PF-escape-1SG let’s go look.for **1INCL.PL** it eat **1INCL.PL**
‘ “I’ve escaped! Let’s go look for something to eat.” ’ (Kuiper & Oram 1991: 407)

Harbour (2016: 149–152) provides some interesting data about the inclusive impostering for other grammatical persons. In (116), risk is incurred if any one individual within the group denoted by *us* is spotted. This suggests that the inclusive includes individuals in its meaning.

(116) If anyone spots **us** (=me / =you / =me and you / =him / =me and you and him / ...), they’ll call the police.

He notes that this impostering property holds of inclusives heralding from different number systems. (117) is from Hawaiian, which is a singular-dual-plural language, and (118) is from Kapampangan, a minimal-augmented language.

(117) a. *Inā ‘ike ‘ia* ***kākou,*** *(e) pono ana (lākou) e ho‘opau i ka pōkā*
if see PASS **1INCL.PL** ASP must ASP 3PL ASP CAUS.finish ACC the shell
pahū ‘ana.
explode NMLZ
b. *Inā ‘ike ‘ia* ***kāua,*** *...*
if see PASS **1INCL.DU**
‘If we get seen, they’ll have to stop bombing.’

(118) a. *Nung atin manakit* ***kekata-mu,*** *miligtas ta-mu.*
if AUX see **1INCL.OBL-AUGM** save **1INCL.ABS-AUGM**
‘If someone sees us (you, me, and others), we’ll be saved.’
b. *Nung atin manakit kekata, miligtas kata.*
if AUX see **1INCL.OBL(.MIN)** save **1INCL.ABS(.MIN)**
‘If someone sees us (you and me), we’ll be saved.’

3.3.2.3 A diachronic affinity between inclusive and impersonal pronouns: Indo-Aryan *APNE*

From §2.2.3 in Chapter 2, readers are already familiar with the Sanskrit impersonal form **atman* ‘self’, later *APNE* (my capitalization here indicates abstraction over its exponential variance across individual languages). *APNE* developed into the 2HH pronouns: *e.g.* Bangla *apna*, Assamese *apuni*, Magahi/Bhojpuri *apne*, and Maithili *əpne*. We have shown that the multiple-tiered politeness systems of these Eastern Indo-Aryan languages involved specific and recurring diachronic pathways from singular to non-honorific; plural to honorific; reflexive to high-honorific, showing that recruitment is not random but rather circumscribed precisely to the pool of ϕ-featural values found in single-tiered politeness systems.

However, these are Indo-Aryan languages which did not innovate a clusivity distinction. In the languages which did innovate clusivity, it was *APNE* which developed into the inclusive

(and also the 2HH pronoun) (Chatterji 1926: 848). It is unclear whether the core of *APNE*'s meaning is very-polite or inclusive, but what is important for our purposes here is the specific and recurring diachronic pathway between reflexive and inclusive. While §3.3.2.1 provided synchronic evidence for the affinity between inclusive and fourth person, this section provides evidence of a diachronic kind. We will use some Southern Indo-Aryan languages to demonstrate.

Some languages in the Southern subfamily straightforwardly recruit plurals for honorification, and do not feature repluralization, unlike their Eastern kin. In Gujarati, the 2NH form *tũ* is inherently singular and takes only singular verbal agreement, while the 2H and 2HH forms *təme, ap* are inherently plural and share the same verbal agreement. Gujarati thus provides a familiar paradigm of recruiting plural for politeness. The lower-politeness tier 2H is recruited from 2PL; the higher-politeness tier 2HH is recruited from fourth person.

(119) Gujarati pronouns (Cardona & Suthar 2003: 739-741, 745)

Gujarati	SG	PL
1EXCL	*hũ, -u, -Ø*	*əme, -ie, -ũ*
1INCL	—	*apṅe*
2NH	*tũ*	—
2H	—	*təme*
2HH	—	*ap*

However, notice another form, the inclusive *apṅe* which bears morphological resemblance to the 2HH pronoun *ap*. (The current-day reflexive in Gujarati has completely evolved away from its Sanskrit form, and is now formed with the reflexive base *pot-*.)

In Marwari, there is a similar affinity between the 2HH pronoun *āp* and the inclusive pronoun *āpaṅ*. Marwari is a language whose current-day reflexive forms, *apɳe ap, apro* (Gusain 2004: 24) still bear resemblance to the Sanskrit. Unlike Gujarati, where there is no repluralization, in Marwari the 2H and 2HH forms can be optionally pluralized with *-log* 'people'.

(120) Marwari pronouns and agreement (Magier 1983: 73-80, 102, MASC, FEM)

Marwari	SG	PL	agreement
1EXCL	*mhẽ, -ã̃*	*mhe*	*-ã̃*
1INCL	—	*āpaṅ*	*-ã̃*
2NH	*thã̃*	—	*-e*
2H	*thẽ*	*thẽ-log*	*-o*
2HH	*āp*	*āp-log*	*-o*
reflexive agentive	*apɳe ap*		
reflexive possessive	*apro*		

Recall from Chapter 2 that, in Marathi, the form *apəṅ* has several different uses: as the form for singular or plural polite address, as the (plural) reflexive, and as the inclusive plural. Thus, Marathi *apəṅ* exhibits the same diachronic affinity between inclusive and fourth person that we have been discussing.

(121) Marathi pronouns and agreement (Dhongde & Wali 2009: 48-50)

Marathi	SG	PL
1EXCL	*mi*	*amhi*
1INCL	—	***apəṅ***
2	*tu*	*tumhi*
reflexive empathic	*swətəha*	***apəṅ***

Lastly, within Eastern Indo-Aryan, Maithili is a language which makes clusivity distinctions. Its inclusive plural is built atop *APNE*, having developed into *əpna səb(h)* (adding *səb(h)* 'all') (Yadav 1996: 107-8).

Chapter 4

Privativity Plus

When you assume,
it makes an ASS out of U and ME.

— Modern proverb

The previous chapter established an empirical landscape with the following landmarks:

- Semantic defaults emerge in contexts of epistemic uncertainty, arbitrarity, and mixed quantification, and cross-linguistically, they take the shapes of plural, third person, and inclusive. So do imposters.
- The shapes of semantic defaults converge with the shapes of polite proxies.

Given that the social-anthropological reflex of the recruitment asymmetries for polite pronouns is negative politeness, the inquiries (and claims) that this chapter explores (and propounds) are:

- Why do the shapes of semantic defaults and proxies of politeness converge?

  Claim: Semantic defaults are presuppositionally weak and emerge in contexts where using a person-neutral and/or a number-neutral form is the only way of being most informative. Proxies of politeness are presuppositionally weak in exactly the same way, being person-neutral and/or number-neutral. Both are presuppositionally weak by dint of having carved out a less specified subregion of the ϕ-lattice.

- Why can presuppositionally weak forms appear in contexts of politeness?

  Claim: Polite contexts are defined as those contexts where a class of pragmatic maxims, termed Taboos, outrank *Maximize Presupposition!*. The Taboos enforce the use of a less-informative (=presuppositionally weaker, underspecified) form, so that speakers maintain deniability and avoid responsibility for the information given. This is the linguistic manifestation of negative politeness: to avoid being specific by avoiding informative-enough forms.

- What does this teach us about the representation of "polite" pronouns?

  Claim: Polite pronouns are expressed via proxies, which serve as negative politeness strategies by dint of being underspecified. No feature dedicated to politeness is necessary.

In §4.1, I build a privative system of ϕ-features. §4.2 shows that, in conjunction with a local strengthening mechanism `exh(aust)`, the proposed system is able to account for the full set of meanings available in a standard tripartition.

The system is then put to the test, deriving both the shapes of semantic defaults and polite pronouns in §4.3 and §4.4, respectively. The commonality that enables their convergence is that strengthening is inoperative. For semantic defaults to arise, strengthening may occur, but only vacuously, since the contexts enabling their emergence are already weakly specified for person and/or number. For polite pronouns to arise, strengthening must not occur, as it is bled by pragmatic maxims I term Taboos. These Taboos are part of a pragmatic calculus for deriving polite meaning.

Crucially, *contra* previous research, politeness is definitionally a ranking of pragmatic maxims, and grammatical features dedicated to politeness are unnecessary in this account. The cross-linguistic distribution and interpretation of honorifics are explained with two independently established ingredients: semantic markedness and negative politeness. The only innovation here

is creating a pragmatic calculus which determines which form is recruited for politeness. This is the main desideratum of this account: extra-grammatical machinery need not be utilized even though politeness is, intuitively, extra-grammatical.

4.1 Linguistic reflexes of indirectness: underspecification via privative features

After our grand tour of semantic defaults, we return back to the crucial question about politeness: given the lack of a feature dedicated to negative politeness and indirectness, what are the analytical possibilities for grammaticalizing these notions? In other words, what are the linguistic reflexes of the Taboo of Directness, as repeated below?

(1) Taboo of Directness (ToD):
Don't use forms which are direct!

In this section, I lay out a feature system that enables us to capture the asymmetric relationship in meaning between proxy (the feature value doing the approximating) and proxand (the feature value being approximated to). This asymmetry reveals itself both in the shapes of semantic defaults established across §3.2 and §3.3, and in the shapes of (un)attested proxies of politeness in Chapter 2.

Strikingly, these two asymmetries converge: *those feature values which are semantic defaults are precisely those values which are recruited for politeness.* I argue that these two asymmetries converge as a natural consequence of the underlying feature system being asymmetrical in the first place, specifically being *privative*: non-default feature values are restrictor functions (derived via positive specification), while default feature values have vacuous meanings (derived via zero specification).

Zero-specified forms may be later strengthened via the application of the local operator `exh(aust)` (Chierchia, Fox & Spector 2012). This way of representing the ϕ-featural system elaborates on the system sketched out by Bobaljik & Sauerland (2023): exhaustification enriches the set of meanings available to a (privative) feature system. Akin to a silent *only*, `exh` is a sublexical entry which induces local strengthening by excluding stronger alternatives. Given a predicate P and its set of alternatives ALT, the exhaustified predicate expresses the conjunction of P and of the negations of all the members of ALT that are not entailed by P. Following Bobaljik & Sauerland, we adopt Mayr's (2015) predicate-level exhaustification operator. This is what Mayr calls `exh2`, defined in (2), but since only predicate-level exhaustification will feature in the analysis, we will call this `exh` throughout.

(2) $[\![\texttt{exh}_{\text{Alt}}]\!]^w = \lambda P_{<\text{e,t}>} \lambda x \in_{\text{e}} .\ P(x)(w) \wedge \forall Q \in Alt\ .\ \neg Q(x)(w) \vee (\forall x(P(x) \rightarrow Q(x)))$

This strengthening operation occurs at the service of the Maxim of Quantity: "Make your contribution as informative as is required, and no more informative than is required" (Grice 1989), an instance of a more general rule governing cooperative communication.

We understand "being informative" as using the strongest form possible in the context. For example, when cardinality is part of the information provided by the context, the speaker is mandated to use the number-specific form instead of a number-neutral form.

Our investigation will consider three distinct usages of pronouns: default, conventional, and polite. Pragmatically speaking, these usages differ crucially in terms of whether strengthening is operative or not, and if it is inoperative, why. Furthermore, there is an interaction with a

privative system of ϕ-features, where features may be valued positively (positive specification) or vacuously (zero specification). The (in)application of `exh` affects the meanings of zero-specified structures differently as below. (The meanings of LFs containing positively-specified structures stay constant across the three distinct uses.)

- In a pronoun's *conventional* use, strengthening is operative. Zero specifications are strengthened by `exh`, and *Maximize Presupposition!* chooses the form most appropriate in the context.
- In a pronoun's *default* use, strengthening is inoperative due to a lack of sufficient information in the context. *Maximize Presupposition!* is inoperative because speakers are unable to make any assumptions about person and/or number. This means that there is no maximally suitable alternative, and only weak forms may emerge.
- Lastly, a pronoun's *polite* use, strengthening is *also* inoperative, but because speakers are mandated not to make any assumptions about person and/or number by a Taboo (detailed below).

» For both default and polite usages, there is an absence of strengthening. Consequently, only semantically underspecified forms—those which are neutral with regard to person and/or number—may emerge.

What is this Taboo which dictates that strengthening must not occur in contexts of politeness? I term it the Taboo on Presumptuousness (Taboo) (3):

(3) Taboo on Presumptuousness (Taboo): Don't make assumptions about one's interlocutor in the conversation.
   a. Don't make assumptions about the number of interlocutors (ω-Taboo).
   b. Don't make assumptions about who the interlocutors are (π-Taboo).

The underlying representations of Logical Forms (LF) are assumed to be constant across all three uses. However, the question of which LFs may surface in which use boils down to an interaction with the pragmatics; namely, whether *Maximize Presupposition!* may apply. We first establish an ontology of persons and numbers and some preliminary assumptions §4.1.1. Building on this is a proposed privative system of ϕ-features §4.1.2.

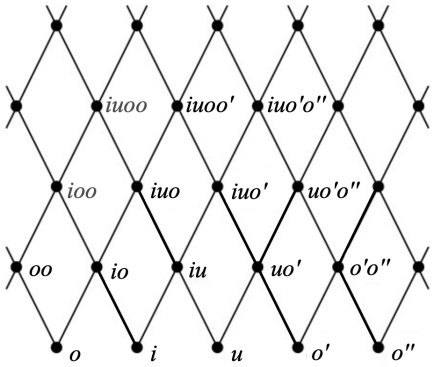
In §4.2, we derive conventional meanings associated with the proposed structures via the application of `exh`. Along the way, we run into the issues of representing impersonals §4.2.1, inclusives §4.2.2, and the the interaction with spellout at PF §4.2.3; all are addressed. It is only in §4.3 that we show how the proposed system derives semantic defaults. Lastly, in §4.4 we finally arrive at the analysis (well, an analysis) of polite pronouns. We will show that different specifications of pragmatic rankings are crucial for deriving, in turn, conventional, default, and polite uses for the set of proposed LFs.

Throughout, there are many choice points made to arrive at a specific implementation. The different choices will be spelt out when we encounter forking paths, but broadly, the general spirit of the analysis remains the same across different implementations. In the most general terms, *strengthening* (implemented here by a local exhaustification operator, which results in stronger presuppositions upstream) variably applies to *meaning-poor* LFs (implemented here via zero specification) to derive the variation in meanings across conventional, default and polite uses.

4.1.1 The ontology and its LF Shadow

Before delving into the technical details, let us establish an assumed ontology of individuals, and how is it represented at LF. First, what may this ontology of individuals look like? Following Link (1983) and others, I assume the graphical representation where individuals are formally organized in a lattice, where atoms represent singulars. Nonsingulars are formed via the join operation $\sqcup$, which corresponds to forming a sum, $a \sqcup b$, out of two distinct objects a and b. I will use ab as a shorthand for $a \sqcup b$. The individuals themselves are assumed to be i, denoting the speaker, u, denoting the addressee, and o, denoting third party (non-speaker and non-addressee). Since the number of o's are infinite, the combinations of i, u, o are also infinite. The person lattice, then, is represented by the Hasse diagram in (4) (with only the first four cardinalities shown). The important ontological assumption is that the combinations of these (non-)participant primitives are ordered in such a way:

(4)



Readers will immediately notice that speaker-denoting i and addressee-denoting u are privileged in the sense that they are unique primitives. An utterance like *I need silence from you* clearly makes available a true plurality of addressees: the request is only successful if all addressees become silent. A plurality of speakers, "choric *we*", is also possible. Thus, the person lattice can also resemble the following[1]:

[1]Harbour (2016: 70) makes a particular point out of the ontological uniquenesses of i and u as follows. I repeat his argument here, although this issue is not consequential for my own account. He acknowledges that choric speaker/hearers are possible, but advocates for an ontology that does not encode them. If the plurality of i's and the plurality of o's are ontologically on par with each other, as (ia-b) supposes, one should be able to replace *he* with *I*. But, this is impossible, as the ungrammaticality of (ii) indicates.

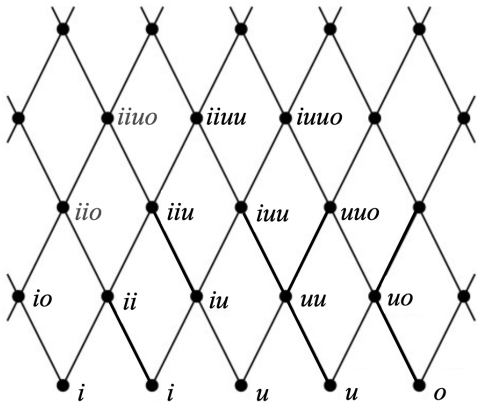
(i) a. We$_{i+i'}$ beseech you for forgiveness for he$_i$ has made hollandaise and I$_{i'}$ have eaten it.
b. They$_{o+o'}$ beseech you for forgiveness for he$_o$ has made hollandaise and he$_{o'}$ has eaten it.

(ii) *We$_{i+i'}$ beseech you for forgiveness for I$_i$ have made hollandaise and I$_{i'}$ have eaten it.

He suggests that a process of metaphoric "othering" occurs in choric contexts: "hearers are egocentric. So, when an addressed group hears a second person plural, each hearer egoistically interprets u as referring to themself and conceives of the others as o's. This exactly parallels the manner in which, in replying with a first person plural, that audience member would consider themself the unique i and the other group members, o's" (Harbour 2016: 71).

Two independent arguments for the ontological uniqueness of i and u can be mounted. Firstly, as discussed and substantiated in Bolbajik (2008), there are no languages which lexicalize the difference between choric *we* ($i+i+i+...$) and non-choric *we* ($i+o+o+...$). The same goes for hearers; there are no languages which lexicalize the difference between choric *you* ($u+u+u+...$) and non-choric *you* ($u+o+o+...$). The second argument comes from the deictic distinctions made within pronominal paradigms. To my knowledge, when deictic distinctions are made in pronouns, such as in some Indo-Aryan languages, they are made only within third persons, such as for Assamese 3SG.MASC: *i* for distal and *xi* for proximal, clearly showing a distinction available for multiple non-speaker-non-hearers. (This deictic distinction is also available for 3SG.FEM, but not maintained in 3PL, Goswami

(5)



For purposes of simplicity and exposition, however, I remain agnostic on this issue of ontological uniqueness, but will assume the uniqueness view. The domain of individuals is limited to the set containing a unique speaker and a unique addressee as in (4), such that the denotation of ϕ is as in (6). It denotes an open-ended lattice with infinite combinations of i, u and o.

(6) $[\![\phi]\!] = \{io*, iuo*, uo*, oo*\}$

A notational note is in order: I adopt the syntax of regular expressions to distinguish between no others, more than one others, and any number of others, as in (7)[2]. Note that, since iu is necessarily a plurality already, its plurality can be notated either as $iuo+$ or $iuo*$.

(7) a. x and no others: x
b. x and one or more others: $xo+$
e.g. $\{io, ioo, iooo, ...\} = \{io+\}$
c. x and any number of others: $xo*$
e.g. $\{i, io, ioo, iooo, ...\} = \{io*\}$

Harbour also assumes that ϕ-features have the following morphosyntactic loci: person features are hosted under the π head, while number features are hosted under the ω head. Thus, the functional spine of the pronoun is as (8): features hosted under the person head π apply to the ϕ-lattice first, then features hosted under the number head ω apply to the resulting lattice. I also adopt his assumption that number features compose after person features[3].

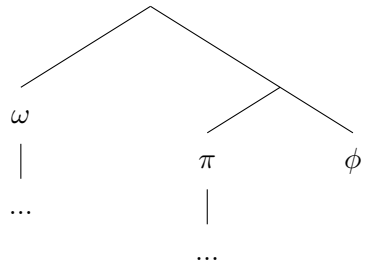
& Tamuli 2003: 420). This suggests that there are no distinctions to be made between multiple speakers and hearers because speakers and hearers are ontologically unique.

[2]Harbour uses the notation whereby "x and any number of others, including zero" are notated x_o. In this notation, subscripted o's are notated with an asterisk as $o*$.

[3]As Harbour (2016: 156-162) notes, this is congruent with Mirror Principle effects, as number morphemes attach farther away from the root than person morphemes do. For instance, consider auxiliary conjugation in Walmatjari (as the auxiliary is the locus of person/number agreement). Taking the auxiliary as the root, we see that person morphemes are closer to the auxiliary than number morphemes are. This is illustrated in (i).

(i) a. *pa- ja- n- ta*
AUX- 1SG- 2- PL
'you all (VERB) me'
b. *ma- n- tarra- nya- lu*
AUX- 2- 1EXCL- DU- PL
'you all (VERB) us two' (Harbour 2016: 159)

(8)



Access to this ontology is mediated by features hosted under these functional heads. I assume two person features hosted under π, [`author`] and [`participant`], and the number feature hosted under ω, [`atomic`]. These are not at all novel features. Following Harbour (2016) (cf. Zwicky 1977), I assume that authors and participants are privileged in the sense that the grammar affords them a dedicated feature, and that hearers are not similarly privileged so that there is no corresponding feature [`addressee`]. Meanwhile, atomicity as a primitive is used widely to derive grammatical phenomena such as the mass-count distinction, collectivity vs. distributivity, amongst many others. Like Harbour, I do not assume that features are themselves organized within feature geometries (Harley & Ritter 2002) or feature hierarchies (Noyer 1992).

The crucial difference from Harbour (2016) is the nature of the feature system. I assume a privative feature system (where features can only take the values of + or 0), whereas he assumes a bivalent feature system (with possible values being instead + and −)[4] I also assume that all features apply to the ϕ-featural lattice via Predicate Modification, which is adapted to apply to sets. Given a node A with daughters B and C, the denotation of A is as in (9).

(9) $[\![\text{A}]\!] = \lambda x_e \,.\, x \in [\![\text{B}]\!] \cap [\![\text{C}]\!]$

When positive feature values [+`F`] are applied to a set, an intersective subset of that set is yielded (10a). When zero feature values F are applied to a set, application is vacuous and the same set is yielded (10b). In essence, applying [+`F`] is to restrict the ϕ-space, while applying [0`F`] is vacuous.

(10) Privative feature action
a. $[\texttt{+F}](S) = \lambda x_e \,.\, x \in [\![\text{F}]\!] \cap [\![\text{S}]\!]$
b. $[\texttt{0F}](S) = S$

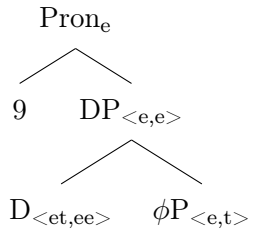
The privative setup, of course, is meant to account for the existence of semantic defaults, for inclusive readings of plural and of third person, established in §3.2. As we shall see later, in scenarios where lack of competition pertains, zero-specified items will end up with weaker meanings, characteristic of semantic defaults and (I argue) polite pronouns.

Harbour (2016: 77) suggests the possibility of presuppositionalizing the contribution of these features, situating the analysis of pronouns in the tradition where ϕ-features are taken to have a presuppositional semantics (Heim 2008; Sauerland 2009), so that the denotational contribution of any pronoun is that of an identity function on individuals (e.g. Elbourne 2001). I take up this suggestion. Traveling to the very top of the functional spine of the pronoun, I assume a high presuppositionalizing head D that applies after all features have applied. It takes a predicate P of individuals and yields a new predicate of individuals that is an identity function which is only

[4] While Harbour acknowledges that all and only the attested partitions can be derived if feature action were to be privative instead of bivalent, he shows that it leads to descriptive inadequacy concerning certain morphological partitions. However, the task here is to derive semantic defaults, instead of morphological partitions, so I will continue with a privative feature system.

defined for individuals that the original predicate P is true of. (For an analogous mechanism for presuppositionalization, see Bobaljik & Sauerland 2023: 84).

(11)



a. $[\![\text{D}]\!] = \lambda P\ .\ (\lambda x : P(x).\ x)$
b. $[\![\text{DP}]\!] = \lambda x_e : x \in P\ .\ x$
c. $[\![\text{Pron}]\!]^g = g(9)$, defined only if $P(x) = 1$

Note that reference to an individual is still accomplished by the index: ϕ-features simply contribute constraints on reference. Presuppositionalization will allow *Maximize Presupposition!* to evaluate LFs as presuppositional alternatives against features of the context. We will see that this is an important component in the impending analysis, where it is the successful application of *Maximize Presupposition!* which derives conventional meanings.

4.1.2 Privative ϕ-featural action on the lattice

In this section, I establish and defend a privative feature system for person and number which captures the facts established above about semantic defaulthood. Privativity will allow us to account for the two asymmetries that we have seen: the recruitment asymmetries found in the typology of polite pronouns from Chapter 2, and the semantic defaulthood asymmetries from §3.2.

The proposal is an extension and revision of Bobaljik & Sauerland's (2023) account of person, where the main innovation is to use a combination of privative features and a local strengthening mechanism (specifically, exhaustification) *within* ϕP to derive the relevant meanings. Here, I flesh out their proposal to also include a privative account for number. However, as §4.2.2 later shows, I diverge significantly from Bobaljik & Sauerland (2023) with regards to the clusivity distinction, the phenomenon which inspires the bulk of their discussion. The difference is that the present account makes novel claims for the representations of impersonals and inclusives, supported by observations on the use conditions on—not morphological exponents of—these categories.

4.1.2.1 Person: Restricting the ϕ-lattice with π features

Let's move on. That snowed-on hanky is someone's but not ours.

— Mother to child, one Massachusetts winter, probably

The denotations of [`auth(or)`] and [`part(icipant)`] are given in (12)-(13). When applied to some set S, their positive specifications restrict S to a subset of S whose members are pluralities that contain i. (the author lattice, in the case of [+`auth`], (12a)) or to a subset of S members are pluralities that contain i or u (the participant lattice, in the case of [+`part`], (13a)). Their zero specifications in (12b) and (13b), on the other hand, apply vacuously, having no effect on S whatsoever, restrictive or otherwise.

(12) a. $[\![[+\texttt{auth}]]\!] = \lambda x_e \,.\, i \sqsubseteq x$
b. $[\![[0\texttt{auth}]]\!] = \lambda x_e \,.\, x$

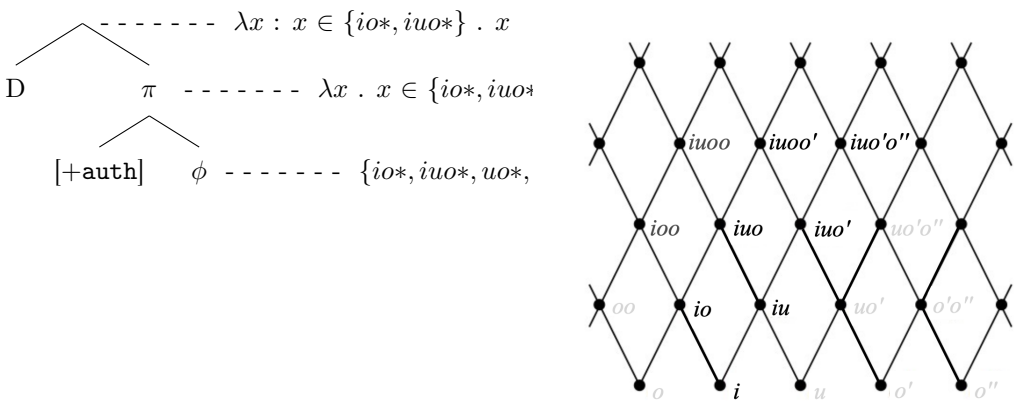
(13) a. $[\![[+\texttt{part}]]\!] = \lambda x_e \,.\, i \sqsubseteq x \vee u \sqsubseteq x$
b. $[\![[0\texttt{part}]]\!] = \lambda x_e \,.\, x$

Importantly, this is a privative system, where values may be specified + or 0, and [+F] entails [0F]. (The corresponding features in Bobaljik & Sauerland (2023) are simply AUTHOR and PARTICIPANT. Indicating the equivalent of positive specification is the presence of either feature in an LF, and indicating the equivalent of zero specification is their absence.) Their effects are not complementary, as opposed to features in a bivalent system, where the values are + or –[5].

Let us now illustrate positive and zero specification of each feature to the ϕ-lattice from (4). Positive feature specifications apply to restrict the lattice, while zero specifications apply vacuously and have no restrictive effect whatsoever.

The strongest person restriction on the ϕ-lattice is delivered by the positive specification [+auth] (which entails [+part]). When [+auth] applies to it, it picks out all and only those set members which contain i. This is indicated by the absence of greying in (14); graphically, this results in an egocentric beam. Recall that, notationally, $\{io*\}$ indicates the set denoting i and any number of others ($\{i, io, ioo, iooo, ...\}$).

(14) The author lattice



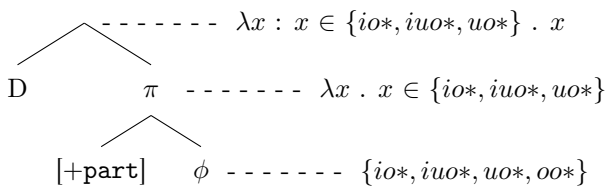
[5] As comparison, in Harbour's bivalent system, the feature action of [author] is as follows. First, [author] as a feature simply denotes the author lattice, which consists only of i (i). Second, positive specification takes every possible pairing of elements in two lattices and joins them together, so that the application of [+author] to the ϕ-lattice is as in (iia), while negative specification cumulatively subtracts every element of one lattice from every element of another lattice as in (iib). Note the difference in notation: Harbour's subscripted o's correspond to asterisked o here.

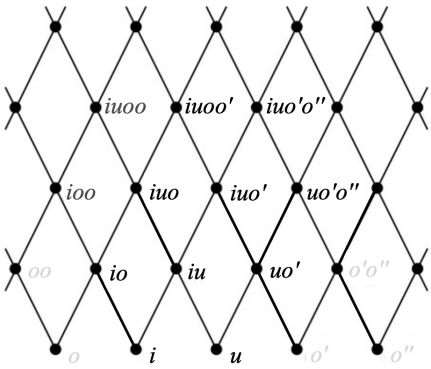
(i) [author] = $\{i\}$

(ii) a. [+author](ϕ)
$= \{i_o, iu_o, u_o, o_o\} + \{i\}$
$= \{ii_o, iiu_o, iu_o, io_o\}$
$= \{i_o, iu_o\}$
b. [-author](ϕ)
$= \{i_o, iu_o, u_o, o_o\} \setminus \{i\}$
$= \{i_o \setminus i, iu_o \setminus i, u_o \setminus i, io_o \setminus i\}$
$= \{\emptyset, o_o, u_o, u_o, o_o\}$
$= \{\emptyset, o_o, u_o\}$

A slightly weaker restriction is delivered by [+`part`]. When [+`part`] applies to the ϕ-lattice, it picks out all and only those set members which contain i or u, resulting in a wider beam, as in (15b). Notice that [+`auth`] entails [+`part`].

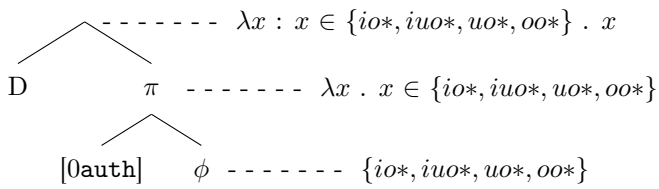
(15) The participant lattice

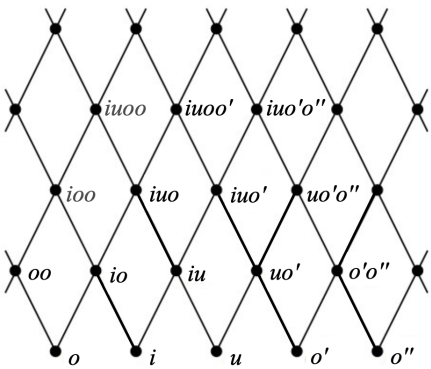




When either [0`auth`] or [0`part`] apply, the ϕ-lattice itself is returned without modification, as in (16c). This delivers the weakest (=no) restriction on the ϕ-lattice. Notice that [0`part`] entails [0`auth`].

(16) The ϕ-lattice, returned





4.1.2.2 Number: Restricting the ϕ-lattice with ω features

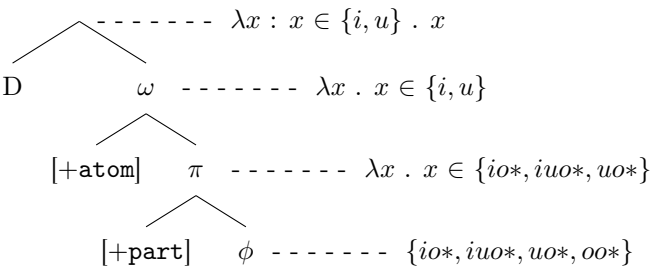
(17) shows the feature specifications of [`atomic`]. Its positive value [+`atomic`] is a function which restricts the lattice to its atomic layer (17a), while its unspecified value [0`atomic`] is an identity function (17b).

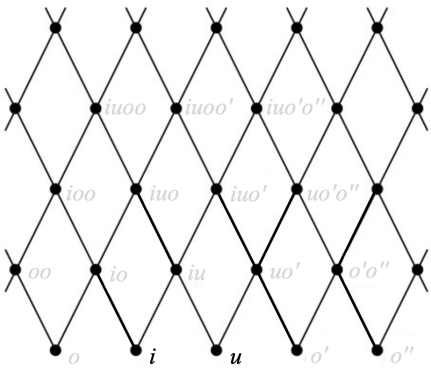
(17) a. $[\![[+\texttt{atomic}]]\!] = \lambda x_e \,.\, \text{atomic}(x)$
b. $[\![[0\texttt{atomic}]]\!] = \lambda x_e \,.\, x$

This asymmetrical conception of number is non-controversial, following much of previous work (Sauerland 2003, 2008; Sauerland, Anderssen, & Yatsushiro 2005), although implementations may differ.

As mentioned, I assume that number features apply after person features. Thus, it is most useful to illustrate the feature action of [`atomic`] to a lattice already specified for person. We can illustrate this with the [+`part`]-specified lattice, the participant lattice, from (15). When [+`atomic`] applies to it, it picks out all and only those participant members which are singularities, i.e. the speaker i and the hearer u, as in (18).

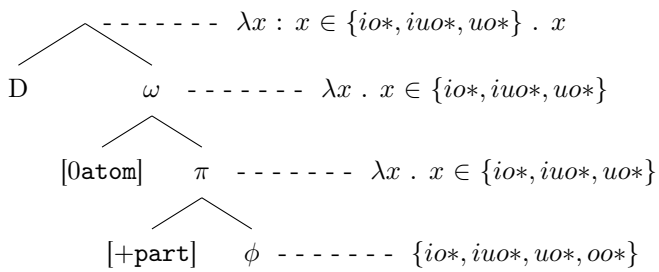
(18) The singular participant lattice

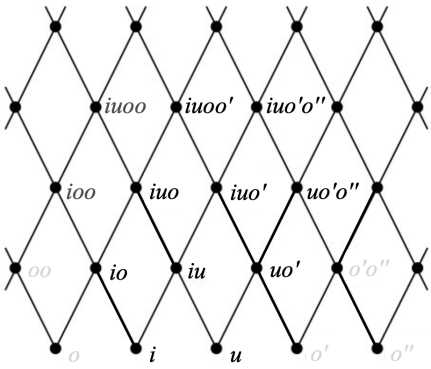




When the [0`atomic`] counterpart applies, the [+`part`]-specified lattice is returned without modification (19).

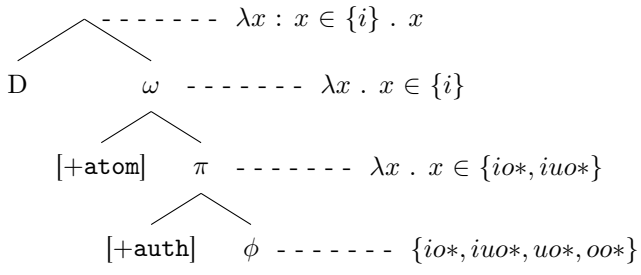
(19) The participant lattice, returned

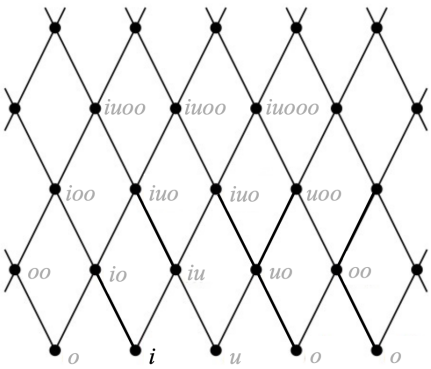




For completeness' sake, we can also illustrate the effect of [atomic] on a [+auth]-specified lattice, the author lattice. When [+atomic] applies to it, it picks out all and only those author members which are singularities, i.e. just i, as in (20).

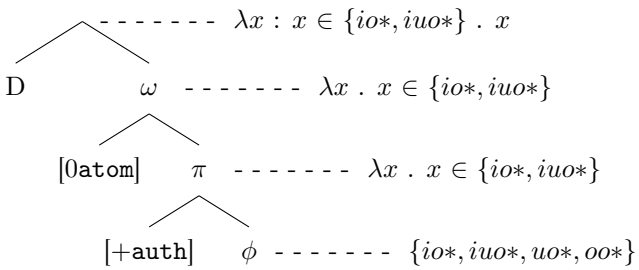
(20) The singular author lattice

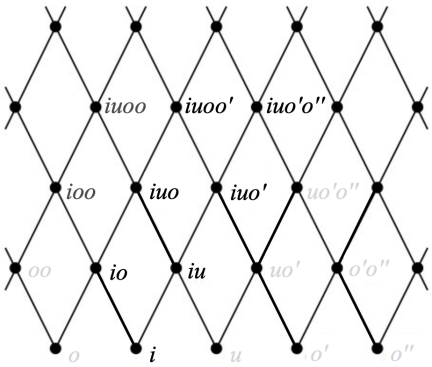




When [0atomic] applies, the author lattice is returned without modification (21).

(21) The author lattice, returned





Those who might be surprised to see plural, not singular, as the feature value represented by the lack of specification at LF, please refer to the discussion on (non-)convergence of markednesses in §3.2.1.

Lastly, readers may have noticed that all zero specifications in the present account ([0`author`], [0`participant`], [0`atomic`]) are synonymously defined as identity functions. To avoid invoking the same meaning three different times, one may prefer a single, more general, definition of zero-specification as an all-purpose identity function that is operative in the absence of positive specification. Making this alteration will have not affect the proposal; however, I will keep the current analytical setup for purposes of exposition, in order to make clear which restriction(s) are not placed on the ϕ-lattice.

4.2 Conventional usages of pronouns: Further restricting the ϕ-lattice via strengthening

Our machinery so far is pitifully lacking: it only generates a set of LFs which correspond to sets containing the author (14), sets containing at least one participant and possibly others (15), and the entire ϕ-lattice unmodified (16). The additional interaction with number features does little to remedy this drastic under-generation, generating a set of LFs which correspond to sets of singular participants (18), of number-neutral participants (19), of singular authors (20), and of sets containing the author (21). Their presuppositions are collated below.

(22) a. $\lambda x : x \in \{io*, iuo*\} \ . \ x$ (14)
b. $\lambda x : x \in \{io*, iuo*, uo*\} \ . \ x$ (15)
c. $\lambda x : x \in \{io*, iuo*, uo*, oo*\} \ . \ x$ (16)
d. $\lambda x : x \in \{i, u\} \ . \ x$ (18)
e. $\lambda x : x \in \{io*, iuo*, uo*\} \ . \ x$ (19)
f. $\lambda x : x \in \{i\} \ . \ x$ (20)

g. $\lambda x : x \in \{io*, iuo*\} \ . \ x$ (21)

This is a far cry from the conventional meanings of pronouns in a standard tripartition, which correspond to the following sets. So far, we have only derived (23a) as (20). Recall that, notationally, $\{xo+\}$ indicates the set denoting x and $\geq$1 number of others (e.g. $\{io+\} = \{io, ioo, iooo, ...\}$).

(23) a. $1\text{SG} = \lambda x : x \in \{i\} \ . \ x$
b. $1\text{PL} = \lambda x : x \in \{io+\} \ . \ x$
c. $2\text{SG} = \lambda x : x \in \{u\} \ . \ x$
d. $2\text{PL} = \lambda x : x \in \{uo+\} \ . \ x$
e. $3\text{SG} = \lambda x : x \in \{o\} \ . \ x$
f. $3\text{PL} = \lambda x : x \in \{oo+\} \ . \ x$

There are two general under-generation issues in the system sketched out so far. First, there is no way to derive sets which contain only the addressee to the exclusion of the author (corresponding to second person) or sets which contain only non-participants to the exclusion of the author and addressee (corresponding to third person). Second, there is no way of deriving honest-to-God pluralities (as opposed to number-neutral sets). We will use a local strengthening operator, `exh` ((2), repeated below), to generate a set of exhaustified LFs and hence a richer set of meanings.

(2) $[\![\texttt{exh}_{\text{Alt}}]\!]^w = \lambda P_{<\text{e,t}>} \lambda x \in_{\text{e}} \ . \ P(x)(w) \wedge \forall Q \in Alt \ . \ \neg Q(x)(w) \vee (\forall x(P(x) \rightarrow Q(x)))$

In what follows, I illustrate derivations using set notation without the accompanying lattice diagrams.

First, to the task of deriving conventional second and third persons, which will involve the application of `exh` to a zero-specified feature, with the alternative set comprising both positive and zero specifications of that feature. This system is essentially equivalent to Bobaljik & Sauerland (2023), although their feature system is defined by the presence *v.* absence of a feature, instead of the positive *v.* zero specification of a feature.

We have seen that structures specified with [+`auth`] feature the strongest restriction on the ϕ-lattice, structures specified with [+`part`] feature a slightly weaker restriction on the ϕ-lattice, and structures specified with [0`auth`] or [0`part`] feature no restriction on the ϕ-lattice whatsoever. These featural specifications are repeated from (12)-(13) below; when applied to the ϕ-lattice, they have the restrictions on the reference of x as in (24).

(12) a. $[\![[+\texttt{auth}]]\!] = \lambda x_e \ . \ i \sqsubseteq x$
b. $[\![[0\texttt{auth}]]\!] = \lambda x_e \ . \ x$

(13) a. $[\![[+\texttt{part}]]\!] = \lambda x_e \ . \ i \sqsubseteq x \vee u \sqsubseteq x$
b. $[\![[0\texttt{part}]]\!] = \lambda x_e \ . \ x$

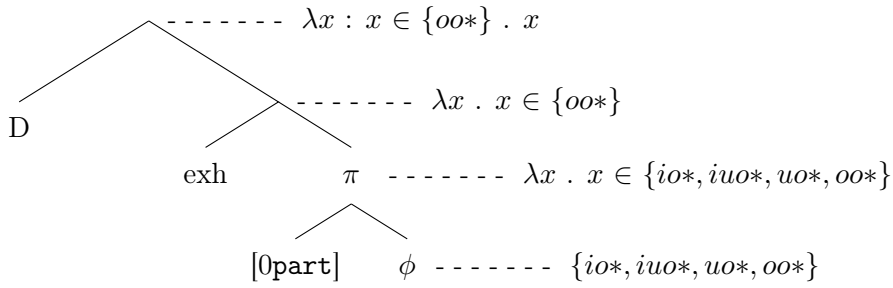
(24) a. $[\![[+\texttt{auth}](\phi)]\!] = \lambda x \ . \ x \in \{io*, iuo*\}$
b. $[\![[+\texttt{part}](\phi)]\!] = \lambda x \ . \ x \in \{io*, iuo*, uo*\}$
c. $[\![[0\texttt{auth}](\phi)]\!] = [\![[0\texttt{part}](\phi)]\!] = \lambda x \ . \ x \in \{io*, iuo*, uo*, oo*\}$

§3.2.2 showed that third person is a semantic default, being person-neutral. So, whence the third person's conventional meaning of "not-participant"? When `exh` applies to [0`part`], it excludes the stronger alternative [+`part`]. A LF with exhaustified [0`part`] is only true of those individuals in the ϕ-lattice (as [0`part`] mandates) but not those individuals in the participant lattice (as the negation of [+`part`] mandates). This is as in (25).

(25) Third persons

a. $[\![\texttt{exh}_{(\texttt{[0part]},\ \texttt{[+part]})}(\texttt{[0part]}(x))]\!] = 1$ iff $x \in \{oo*\}$, 0 otherwise

b.



This correctly derives the conventional meaning of the third person, "not-participant". Hence, [0part] LFs are born person-neutral. They may obtain this conventional meaning via exhaustification, *not* from any kind of hardcoded non-participanthood as in (26). (Regardless of whether there is strengthening or not, at PF, [0part] forms are spelt out as the morphological third person.)

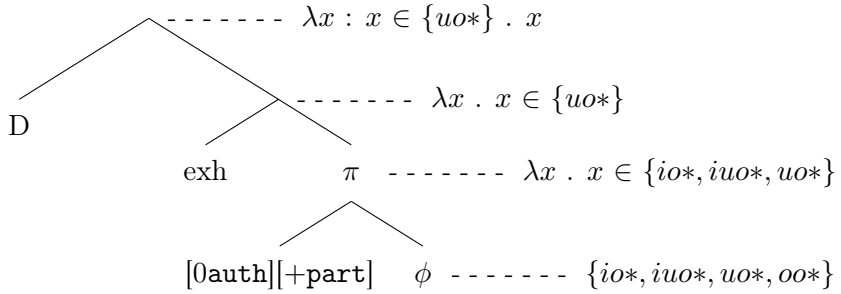
(26) $[\![\texttt{[0part]}]\!] \neq \lambda x_e\ .\ \neg\text{participant}(x)$

A similar process derives the conventional meaning of second person, "participant-but-not-author". When exh applies to [0auth][+part], it excludes the stronger alternative [+auth]. A LF with exhaustified [0auth][+part] is only true of those individuals in the participant lattice (as [0auth][+part] mandates) but not those individuals in the author lattice (as the negation of [+auth] mandates). This is as in (27).

(27) Second persons

a. $[\![\texttt{exh}_{(\texttt{[0auth][+part]},\ \texttt{[+auth]})}(\texttt{[0auth][+part]}(x))]\!] = 1$ iff $x \in \{uo*\}$, 0 otherwise

b.



Hence, we have illustrated two possible ways to restrict the lattice, where the meanings of positively-specified structures and zero-specified structures are restricted in different ways. The former's meanings are restricted by ϕ-features themselves. Meanwhile, the latter's meanings are restricted by exh, whose application to zero-specified forms results in stronger meanings via negation of the meaning of its positively-specified counterparts.

Now, to the task of deriving honest-to-God pluralities. The proposal for an asymmetrical semantics for number is repeated from (17). [+atomic] features a restriction to the atomic layer of the ϕ-lattice, while [0atomic] features no restriction whatsoever.

(17) a. $[\![\texttt{[+atomic]}]\!] = \lambda x_e\ .\ \text{atomic}(x)$

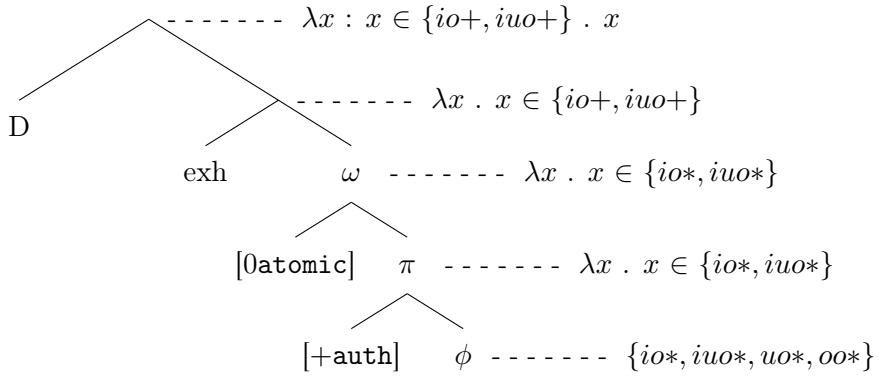
b. $[\![\texttt{[0atomic]}]\!] = \lambda x_e\ .\ x$

Data from §3.2.3 showed that plural is a semantic default, being number-neutral. So, whence the plural's conventional meaning of "more-than-one", or "not-singular"? When `exh` applies to [0`atomic`], it excludes the stronger alternative [+`atomic`]. A LF with exhaustified [0`atomic`] is only true of those individuals in the ϕ-lattice (as [0`atomic`] mandates) but not those individuals in the atomic layer (as the negation of [+`atomic`] mandates). (28) illustrates the derivation of the first person plural, taking an example with [+`author`] specified at π. (Recall that $iuo*$ and $iuo+$ are merely notational variants of the same lattice subregion.)

(28) First person plurals

a. $[\![\texttt{exh}_{([0\texttt{atom}][+\texttt{auth}],\ [+\texttt{atom}][+\texttt{auth}])}([0\texttt{atom}][+\texttt{auth}](x))]\!] = 1$ iff $x \in \{io+, iuo+\}$, 0 otherwise

b.

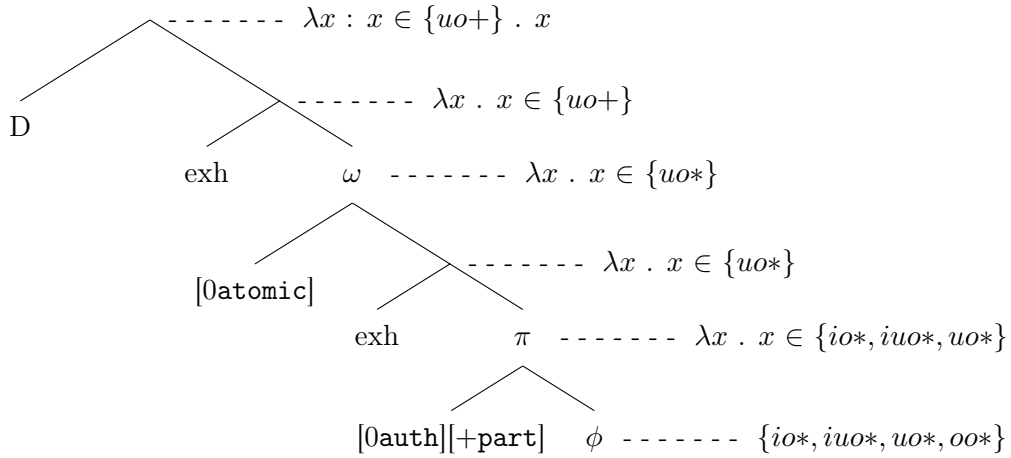


Hence, [0`atomic`] forms are born number-neutral, and may obtain their conventional meaning of "not-singular" via `exh`-mediated competition, but *not* from any kind of hardcoded multiplicity as in (29). (Regardless of whether there is strengthening or not, at PF, [0`atomic`] forms are spelt out as the morphological plural.)

(29) $[\![[0\texttt{atomic}]]\!] \neq \lambda x_e\ .\ \neg\text{atomic}(x)$

(30) illustrates the derivation of the second person plural. There are two steps of exhaustification. At π, `exh` excludes the author lattice from the participant lattice, leaving behind the addressee lattice; at ω, `exh` excludes the atomic layer from the addressee lattice.

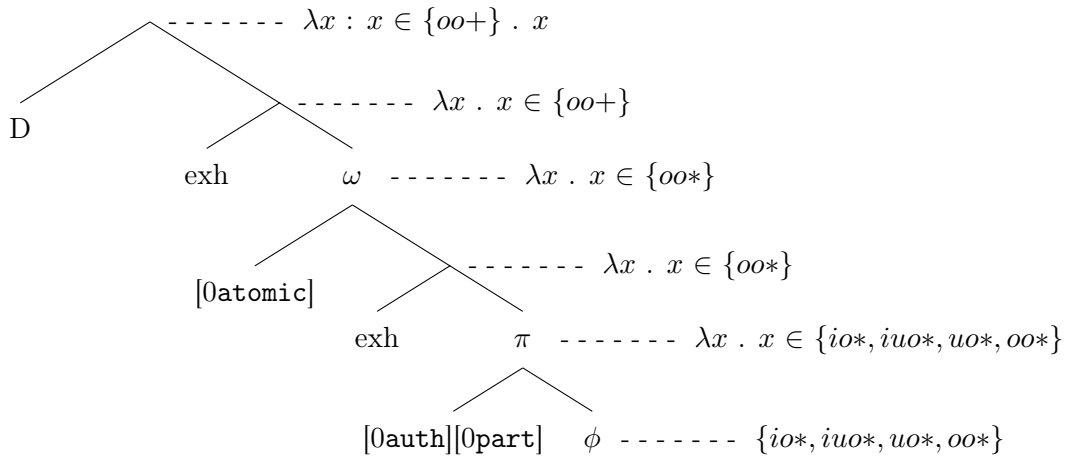
(30) Second person plurals



(31) illustrates the derivation of the third person plural, also via two steps of exhaustification.

At π, `exh` excludes the participant lattice from the ϕ-lattice, leaving behind the non-participant lattice; at ω, `exh` excludes the atomic layer from the non-participant lattice.

(31) Third person plurals



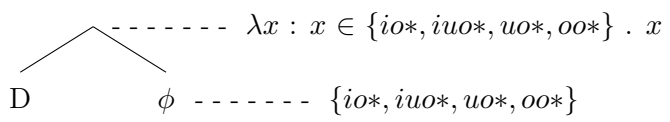
We have now addressed the undergeneration problem outlined at the beginning of this section, deriving the second and third persons, plus plurals for all persons. This richer set of meanings is generated via two possible routes to restricting the ϕ-lattice: positive specification or exhaustification. If positive specifications are involved, they directly restrict the ϕ-lattice, eliminating elements which do not contain i (in the case of [+`auth`]), eliminating elements which do not contain i or u (in the case of [+`part`]), or eliminating non-atomic elements (in the case of [+`atomic`]). If zero specifications are involved, their application to the lattice is vacuous and any restrictive effect they have on the ϕ-lattice only comes about via exhaustification. Exhaustification as applied to a zero specification[6] restricts its denotation by excluding the subregion picked out by its positively-specified counterpart, thereby strengthening its meaning.

4.2.1 Impersonals: Wanton non-restriction of the ϕ-lattice

Restriction occurs when the person and number heads project in the first place—when these heads carry person or number features. For impersonals, I assume that these projections are lacking. I analyze impersonals as the bare ϕ lattice, which does not project neither the π head for person features, nor the ω head for number features, and merely denotes the set of individuals. This will capture a hallmark of impersonals that we have seen in §3.2.4: its utter number- and person-neutrality.

The impersonal merely denotes that there is *some* entity involved, but it does not pick out any particular one from the ontology. (To capture its other hallmark, its animacy requirement, we brute-force assume that the ϕ-lattice is the domain of animates, such that an inanimate pronoun like English *it* cannot refer to the o's contained within.) The impersonal pronoun could refer to any member within this (very large) set, depending on the value of the referential index upstream.

(32)



This structure is unique in the sense that strengthening cannot apply to it; it always picks

[6] In principle, `exh` can also apply to a positive specification, but this would be vacuous.

out the entire ϕ-lattice. Featuring absolutely no specification at all (at either π or ω), sentences containing the impersonal are impervious to `exh`-mediated competition, and there is no alternative set associated with the impersonal.

The precise implementation for deriving its conventional meaning ("some entity within the ϕ-lattice") does not matter for my purposes here, as I only require the impersonal to be a highly impoverished alternative. The interested reader is referred to several analyses on the market. Many scholars (Moltmann 2006; Sigurðsson & Egerland 2009; Zobel 2024, *a.o.*) treat impersonals as featurally impoverished variables bound by a generic operator Gen higher in the structure; Hall (2019) argues for an additional choice function which picks out contextually determined subsets of the ϕ-lattice.

This proposal captures the fact that impersonals display a one-way quacking relation: impersonals never quack like anyone else, and only other ϕ-featural values quack like them. So, why do impersonals never quack like any other? This is because the impersonal carries such a weak meaning that a speaker would only use it when when there is no information whatsoever about person or number assumed in the context. Further, why do other ϕ-featural values like 3PL quack like the impersonal? This is because in scenarios where there is no information about both person and number, the pronoun contributes the same, very weak meaning as the impersonal does, as in (32).

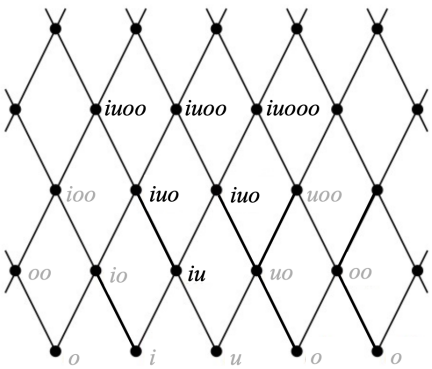
4.2.2 Inclusives

In this section, I propose a novel LF for the inclusive plural, starting from the observation that it may be used impersonally, as §3.3.2 showed. I sketch how its impersonal interpretation must differ from its conventional interpretation. Existing analyses of the inclusive only account for its conventional interpretation, and I propose a LF for the inclusive plural where the π projection for person is crucially absent.

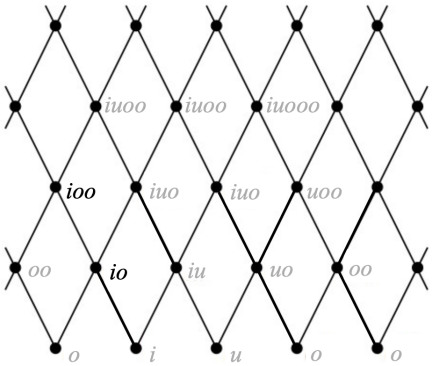
4.2.2.1 The inclusive: impersonal and conventional interpretations

We turn to the conventional interpretation of the inclusive first. In languages which make a clusivity distinction, the first person inclusive plural refers to a group composed of both speaker and addressee, and possibly one or several third persons, as in (33). In contrast, the *exclusive plural* refers to a group composed of speaker and one or several third persons—without the addressee, as in (34). Both are plurals.

(33) The conventional inclusive's partition of the ϕ-lattice

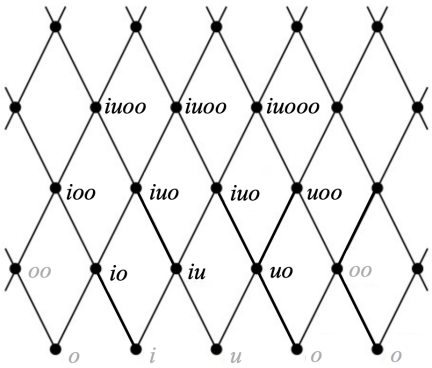


(34) The exclusive's partition of the ϕ-lattice



English does not make a clusivity distinction, so that its first person plural may refer to either group (myself, one or several third persons, with or without you; also termed *general first person*). This is merely the participant lattice, the conjunction of (33) and (34), as in (35).

(35) The general first person's partition of the ϕ-lattice

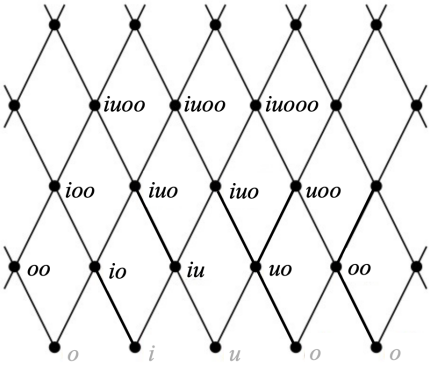


However, an analysis of the inclusive must capture the fact that it is a semantic default of sorts, being able to appear in contexts of arbitrarity in some languages lacking dedicated impersonsal pronouns. This is the inclusive's impersonal interpretation. Recall the novel observation from §3.3.2.1 that, in languages like Fehan Tetun, it is the first person inclusive plural that appears in prescriptive statements in (91), repeated below. In this context especially, the conventional interpretation does not pertain: the speaker is describing women's work to a male addressee.

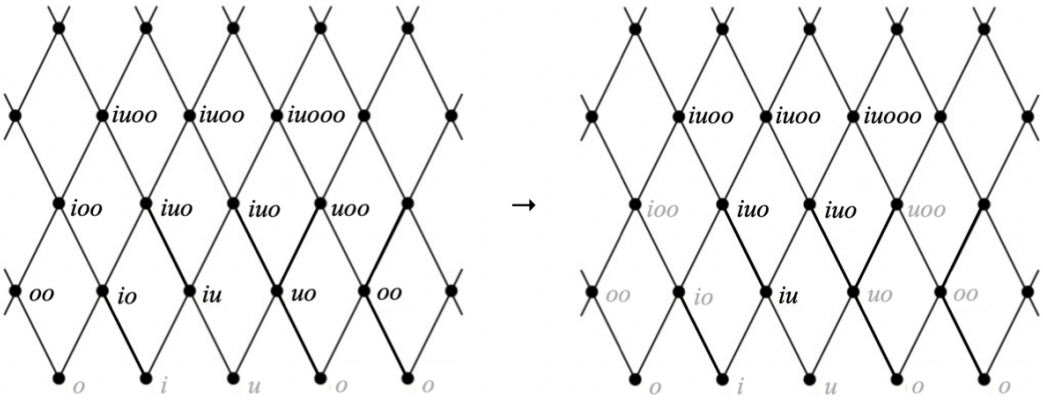
(91) ***Ita*** *feto la'o ha'i.* ***Ita*** *iha uma karian.*
1INCL.PL woman walk not **1INCL.PL** LOC house work
'We (generic) women don't go out. We (generic) work in the house.' (said to a male)
(van Klinken 1999: 112-3)

Based on this, I argue that the inclusive must have a person-neutral interpretation at least at some point in its derivation because it is unspecified for grammatical persons; that allows the inclusive to be used for prescriptions not directed at specific interlocutors. Its impersonal interpretation, schematized in (36), must be differentiated from its conventional interpretation somehow, which was schematized in (33). I propose that the impersonal inclusive carves out a person-neutral space, ranging over pluralities of the speaker, addressee, and third persons.

(36) The impersonal inclusive's partition of the ϕ-lattice



Crucially, since it ranges over the pluralities of all persons, it is person-neutral and hence a semantic default. Granted, the generality of a prescription could originate from a higher generic operator Gen over an individual variable instead of from plural reference. However, we pursue the plural option for now, as our goal is to capture the semantic defaulthood of the inclusive on par with the other defaults. The impersonal inclusive, then, should first have a LF which closely resembles that of the impersonal, as proposed. An additional step of strengthening would derive the conventional inclusive's partition from the impersonal inclusive's partition, winnowing down the ϕ-space from (36) to (33), as below.



We must mention, however, that equating the inclusive as a person default, however, flies in the face of existing ϕ-featural theories. This is because the inclusive plural is derived via multiple (positive) specifications, as §4.2.2.2 will show.

4.2.2.2 Existing analyses of the inclusive

Bobaljik & Sauerland's (2023) first step in deriving the inclusive is the exhaustification of PARTICIPANT to yield second person, as in (37a). Secondly, non-Boolean conjunction[7] adds this result with a subset containing only members with authors ([+`auth`] applied to ϕ), as in (37b)[8]. The tree in (37c) illustrates this proposal (with the present ϕ-featural system). A local operator `exh` is responsible for strengthening. The final meaning $\lambda x\ .\ x \in \{iuo*\}$ is precisely the inclusive's partition of the person space, as was schematized in (33).

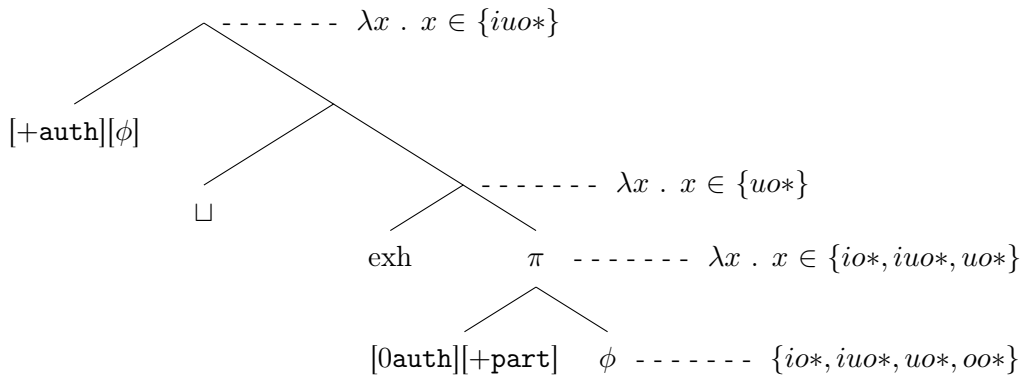
[7]While Boolean conjunction is intersective, non-Boolean conjunction is cumulative (see Schmitt 2020).

[8]They posit the availability of non-Boolean conjunction only for languages that make a clusivity distinction.

(37) a. $\lambda x\ .\ x \in \{io*, iuo*, uo*\}\ \wedge \notin \{io*, iuo*\}$
$= \lambda x\ .\ x \in \{uo*\}$

b. $\lambda x\ .\ x \in \{uo*\}\ \sqcup \{io*, iuo*\}$
$= \lambda x\ .\ x \in \{iu, iuo\}$
$= \lambda x\ .\ x \in \{iuo*\}$

c.



Hence, the inclusive contains a second person core obtained via exhaustification, atop of which the first person layer is "added back" via non-Boolean conjunction. However, the derivation in (37) does not capture the inclusive's ability to approximate an impersonal, because there is no step of the derivation in which the structure is person-neutral. It is likely that it is the existence of an operator like GEN atop (37c) derives impersonal meaning, but the authors do not spell this out. We also do not pursue this possibility for now, as we just require that the inclusive is somehow underspecified for some ϕ-featural category.

Another analysis that requires mention is Harbour (2016). The feature system there is bivalent: positive feature action of [+F] on lattice P indicates point-wise addition of F to P, while negative feature action of [−F] on lattice P indicates cumulative subtraction of F from P. The person features are also [author] and [participant], with each denoting the author lattice and the participant lattice, respectively (38).

(38) a. $[\![[\texttt{author}]]\!] = \{i\}$ author lattice
b. $[\![[\texttt{participant}]]\!] = \{i, iu, u\}$ participant lattice

The principle of Lexical Complementarity (L.C.) has a similar effect to strengthening. It induces a partition when the denotations of two feature combinations are in a subset–superset relation. Informally, it states that the feature combination with the larger denotation is used only when the smaller one is inappropriate, "catch[ing] semantic slack". For instance, if there is a structure A which denotes the set in (39a), and a structure B which denotes a subset of this set in (39b), then Lexical Complementarity kicks in to strengthen A's denotation to one which excludes the set members of B (39c)[9].

(39) a. $[\![\text{A}]\!] = \{io*, iuo*, uo*\}$
b. $[\![\text{B}]\!] = \{io*, iuo*\}$
c. $[\![\text{A}]\!]$ after L.C. $= \{uo*\}$

[9] As far as I understand, Lexical Complementarity is not a mechanism sensitive to context: once it applies, it carves out complementary sets of meanings. To account for the asymmetrical meanings of nominal and pronominal number, Harbour (2016: 150-1) appeals to "atomic restitution", whereby something in the context "permits the unpacking of pluralities and the restitution of the atoms that they contain". Since semantic defaulthood was not within the scope of the analysis, the exact mechanism is not fully spelt out, with a choice function being offered as a possibility. Under this approach, the aspect of the utterance context that enables dismantling must at least contain polite contexts, where we see the effects of atomic restitution as well. Furthermore, as the name suggests, atomic restitution does not account for semantic defaults in the category of for person.

To derive a system with a clusivity distinction (a quadripartition 1INCL | 1EXCL | 2 | 3), [`participant`] applies first to the ϕ-lattice, then [`author`]. The derivation of the exclusive is illustrated in (40a), while the inclusive in (40b).

(40) a. [+`author`] ([−`participant`](π))
$= \{\emptyset, oo*\} + \{i\}$
$= \{i, ioo*\}$
$= \{io*\}$

b. [+`author`] ([+`participant`](π))
$= \{io*, iuo*\} + \{i\}$
$= \{io*, iuo*\}$ (via L.C. since this is a superset of (40a))
$= \{iuo*\}$

Here too, the inclusive (40b) necessarily involves strengthening via a competitor (40a) in its derivation, just as in Bobaljik & Sauerland (2023). However, again, there is no stage in the derivation to accommodate a person-neutral interpretation. The pre-strengthened set, $\{io*, iuo*\}$, is not approximate enough to the set denoted by an impersonal, $\{io*, iuo*, uo*, oo*\}$.

4.2.2.3 A small intervention

The failure of systems to derive the semantic defaulthood of the inclusive may indicate the need to reconsider multiple aspects of ϕ-featural theory, such as the privileged status of authors and participants, the kinds of feature specifications that are permissible, or the organization of ontological primitives into lattices. These aspects have traditionally been exclusively informed by *morphological* factors, notably the morphological distinctions, or partitions, made within a paradigm. Its establishment of natural classes, such as author and participant, is informed by which groupings of ontological primitives are targeted together for spellout at PF.

However, this dissertation has emphasized (along with Sauerland and others) that *semantic* factors such as semantic defaulthood should also be used to inform ϕ-featural theory. It asks which groupings of ontological primitives can stand in semantically for which other groupings, and the ranges of meanings different forms can take. Thus, it establishes containment relationships at LF (which do not necessarily dovetail with other relationships at PF), instead of partitions at PF.

Thus, the schematization of paradigms will necessarily be very different depending on whether we approach them from a PF perspective or LF perspective. (41) are subject pronoun paradigms adapted from Harbour (2016: 51), which explicitly frames the "partition problem" from a morphological standpoint. Assuming that the maximal underlying distinction for person is the full panoply of 1EXCL-1INCL-2-3, only 4 types of distinctions are made at spellout (41a) (monopartition is also attested, but not for subject pronoun paradigms). This is the heart of the partition problem: other ways of carving up the space of spellout are possible, but unattested; for instance, the bipartition 1EXCL | 1INCL, 2, 3 or the tripartition 1EXCL | 1INCL, 2 | 3.

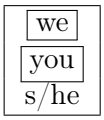
(41)

<table>
<tr><th></th><th>Imonda</th><th>English</th><th>Damin</th><th>Hocąk</th></tr>
<tr><td>1EXCL</td><td>ka</td><td rowspan="2">we</td><td rowspan="2">n!aa</td><td rowspan="3">nee</td></tr>
<tr><td>1INCL</td><td>pəl</td></tr>
<tr><td>2</td><td>ne</td><td>you</td><td rowspan="2">n!uu</td></tr>
<tr><td>3</td><td>ehe</td><td>s/he</td><td>'ee</td></tr>
</table>

From a LF perspective informed by semantic defaults, however, one would annotate the ontological space with containment relationships instead of splitting subspaces from each other. The

English paradigm, for instance, would be represented as (42), capturing the person-neutrality of the third person.

(42)



This teaches us a general lesson about the additional phenomena we can, or should, use to inform ϕ-featural theory. This dissertation shows that LF phenomena, like semantic defaulthood and use conditions, should also be considered alongside PF phenomena, like morphological exponence.

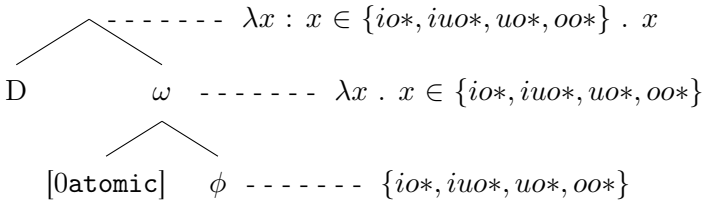
4.2.2.4 A tentative proposal for the inclusive: Two steps of strengthening

A key component of our analysis for the inclusive will depend on making a distinction between the different routes to impersonality established in §3.2.4: plural predication or distributive quantification. I will argue that the " impersonal inclusive" takes the route of plural predication, denoting the non-atomic layer across all persons. This analytical move will put the impersonal inclusive on par with other semantic defaults. Since semantic defaults are uniformly underspecified for some ϕ-featural category, the inclusive is person-neutral.

A tentative proposal for the impersonal inclusive is as follows. I crucially advocate for the non-projection of the locus of person features, the head π. This captures its person-neutrality. The non-projection of this head is equivalent to being zero-specified for both person features, but crucially, it precludes strengthening along the person dimension as π is not projected in the first place. Thus, the impersonal inclusive may only be strengthened for number.

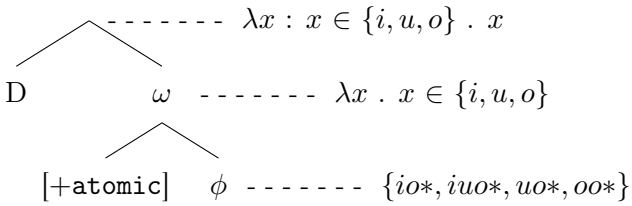
(43) shows its pre-strengthened state: it contributes the same very weak presupposition as the impersonal (that its referent be animate). Happily, this makes inclusives very structurally similar to impersonals, a desideratum we sketched out in §4.2.2.1. This presuppositional similarity explains the possibility of using an inclusive to approximate an impersonal.

(43)



To put it lightly, the "only" problem is that its competitor, a structure with the minimal difference of being specified [+atomic], is not known to be morphologically productive. It would be a person-neutral singular pronoun, as in (44).

(44)



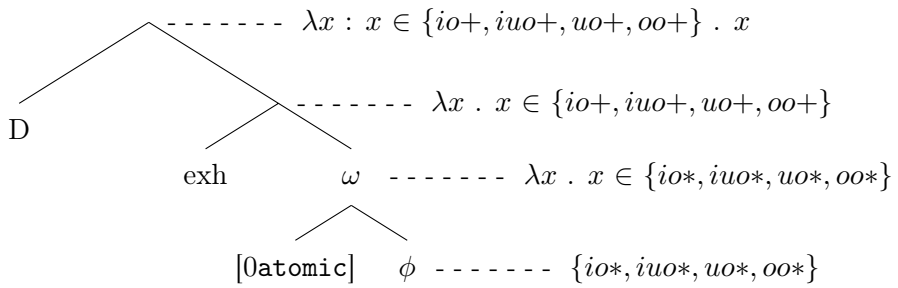
However, such a structure is necessary for the proposed competition as in (45). A LF with

exhaustified [0atomic] is only true of those individuals in the ϕ-lattice (as [0atomic] mandates) but not those individuals in the atomic layer (as the negation of [+atomic] mandates).

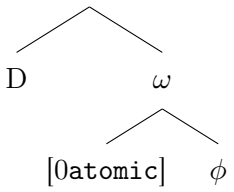
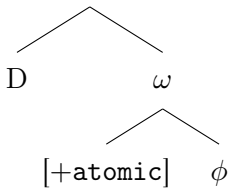
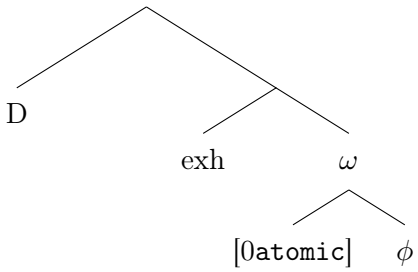
(45) Impersonal inclusives

a. $[\![\texttt{exh}_{([0\texttt{atom}],\ [+\texttt{atom}])}([0\texttt{atom}](x))]\!] = 1$ iff $x \in \{io+, iuo+, uo+, oo+\}$, 0 otherwise

b.



This LF correctly captures the inclusive's plural predication flavor (it is only suitable directed towards strict pluralities) and its suitability for use in contexts of arbitrarity (being person-neutral, it ranges over all the grammatical persons). This step of strengthening along the dimension of number is illustrated below.

① Weak LF	② Stronger alternative	③ Strengthened LF
Individuals $\lambda x : x \in \{io*, iuo*, uo*, oo*\} . x$  	Singular person-neutral $\lambda x : x \in \{i, u, o\} . x$  	Plural person-neutral $\lambda x : x \in \{io+, iuo+, uo+, oo+\} . x$  "not-singular"

However, yet another step of strengthening is required to derive the conventional meaning of the inclusive, which denotes a narrower set: 'me and you and any number of others'. The proposed impersonal-inclusive currently has the denotation in (46a). It must undergo competition with a competitor (46b) to arrive at the conventional inclusive's restriction of the ϕ-lattice in (46c).

(46) a. $\lambda x : x \in \{io+, uo+, iuo+, oo+\} . x$

b. $\lambda x : x \in \{io+, uo+, oo+\} . x$

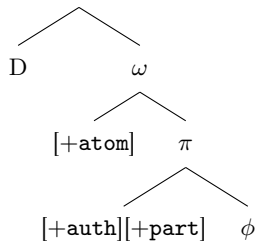
c. $\lambda x : x \in \{iuo+\} . x$

I do not know how to implement this second instance of strengthening to derive the conventional meaning of the inclusive within the current system. Bobaljik & Sauerland's (2023) solution is the non-Boolean conjunction of first persons to a second person core, as detailed above, but I remain agnostic about this possibility and leave it to future work.

4.2.3 PF shadows

All this talk about Logical Form may leave some readers wanting some discussion of the shadow it casts at PF. Here, I collate the LFs established with their corresponding spellouts at PF (in English, but is assumed to map onto any language with a similar system, the so-called standard tripartition). We assume a Nanosyntactic model of spellout (Starke 2002, 2009; Caha 2009), where a single morpheme in the lexicon may correspond to a span of several terminal nodes. Thus, in (47) below, *I* is the spellout for that precise sequence of terminal nodes, and not the spellout for any other sequence.

(47)



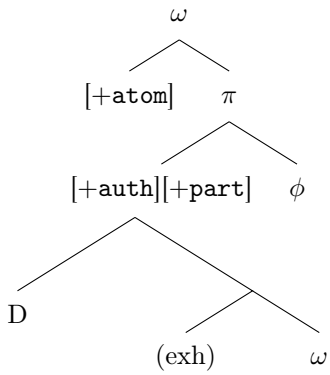
$\Leftrightarrow I$

In contrast, in Distributed Morphology (DM), *feature bundling* must be assumed to handle such a one-to-many relationship between lexical entries and terminal nodes. Thus, the DM picture of spellout for (47) would be as (48). I remain agnostic about the abstract issue of whether feature bundling is possible or not; for differences between Nanosyntax and Distributed Morphology, see Caha (2020).

(48) [+atomic][+auth][+part] $\Leftrightarrow I$

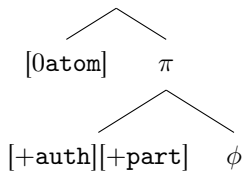
Assuming the Nanosyntactic conception of spellout, then, (49) illustrates the LF-PF correspondences for those structures projecting both functional heads. Recall that, regardless of whether there is strengthening or not (as indicated by bracketed exh), at PF, [0part] correspond to third person, and [0atom] forms to plural.

(49) a.

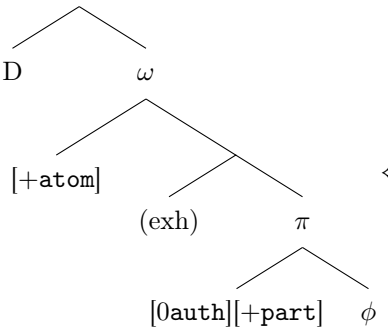


$\Leftrightarrow I$

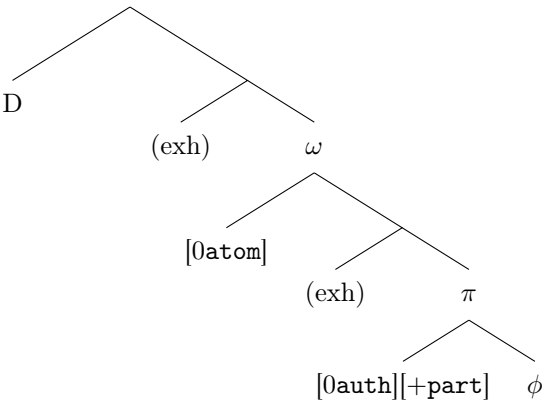
b.



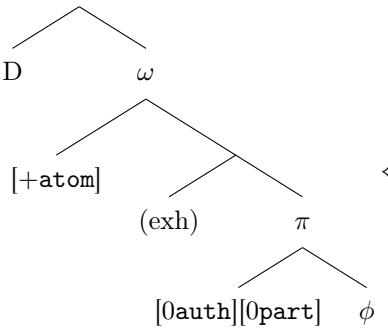
$\Leftrightarrow$ *we*

c. 

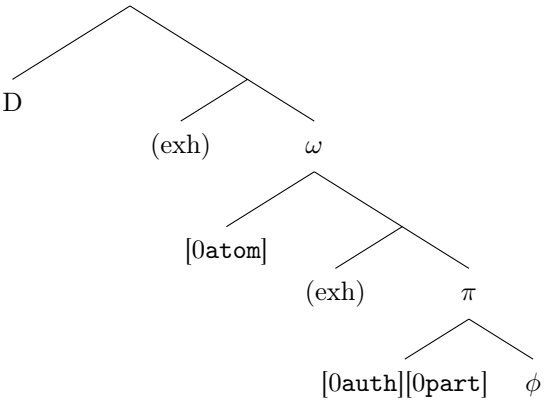
⇔ *you*

d. 

⇔ *you*

e. 

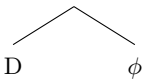
⇔ *he/she*

f. 

⇔ *they*

(50) illustrates the LF-PF correspondences for the more impoverished structures projecting only one functional head; namely, the inclusive from §4.2.2. Recall from §4.2.2.4 that our analysis of the inclusive as in (50a) requires competition wth an [+`atomic`] counterpart which is not known to have any spellout as in (50b).

Lastly, (50) illustrates the LF-PF correspondence for the absolutely structurally impoverished impersonal from §4.2.1. English is a language with a dedicated impersonal form (*one*), but this is not universal.

(50) 

$\Leftrightarrow$ *one*

4.3 Default usages of pronouns: Strengthening has defaulted

So far, we have only considered conventional usages of pronouns, where `exh` applies without a second thought. The posited structures are sometimes in competition with each other, but otherwise live happily, with a rich set of meanings distributed amongst them.

The object of this section is to derive default usages of pronouns, where the crucial difference is that `exh` does not apply. Since there is no competition, zero-specified forms retain their weak meanings. The lack of competition leaves a trail of inequality upstream at DP: positively-specified structures are presuppositionally strong relative to their zero-specified comrades, which are unstrengthened and are presuppositionally weak.

This asymmetry between zero-specified and positively-specified structures will enable us to explain the shape of semantic defaults. We will see that all ϕ-featural values serving as semantic defaults are left unstrengthened, or are unstrengthenable, for some dimension of meaning:

- Plural is the default number *because it is number-neutral when left unstrengthened*;
- Third person is the default person *because it is person-neutral when left unstrengthened*;
- The impersonal inclusive is a default *because it is person-neutral when left unstrengthened*;
- The impersonal is a default *because it is unstrengthenable* and thus always number- and person-neutral.

In contexts where there the speaker makes no assumptions about person and/or number, *Maximize Presupposition!* vacuously applies.

(51) *Maximize Presupposition!*:
Choose the strongest presupposition compatible with what is assumed in the conversation.

In contexts of epistemic uncertainty, the speaker is uncertain about which specific feature value holds. In contexts of arbitrarity, all existing feature values would be too specific. In contexts of mixed quantification, no existing feature value would be specific enough. The common denominator of these contexts is that the neutral form is already the "maximal" feature value that is compatible with what is assumed in the conversation. (And to use a non-neutral form would result in presupposition failure.) After all, there are no ϕ-featural values specialized for epistemic uncertainty, arbitrarity, or mixed quantification.

Let us briefly illustrate the inapplicability of MP! with our lottery scenario from (8). In this scenario of epistemic uncertainty, recall that the third person surfaces.

(8) *Context: Andreas, Bobcat, and Caraway have just bought lottery tickets. Nobody knows who will win. Andreas says to Bobcat:*
   a. #Whoever wins the lottery will be so lucky... it will be **my** ticket outta this sorry-ass town.
   b. #Whoever wins the lottery will be so lucky... it will be **your** ticket outta this sorry-ass town.

c. Whoever wins the lottery will be so lucky... it will be **his** ticket outta this sorry-ass town.

Here, it is not assumed that the winner would be the speaker, addressee, or the third party. MP! cannot enforce the choice of a person-specific form, and only the person-neutral value is suitable. (If the person-specific values, first/second, were used, that would result in presupposition failure.)

I argue that the defining aspect of contexts where semantic defaults emerge is that MP! cannot apply. Based on this, I argue that default forms are weak in meaning. This comes about from two specific happenstances: zero-specification and the non-application of `exh`. The first, zero-specification, we have already seen extensively. The second, non-application of `exh`, is but a minimal difference, and will be illustrated now for default persons. We will see that zero-specified pronouns retain their weak meanings in the absence of strengthening, making them suitable semantic defaults.

The table below illustrates the semantic defaults for the category of person, which are defined by the absence of `exh` at π. [0`auth`][+`part`] LFs retain their weak meaning "participant" (instead of being strengthened to "participant-but-not-author" as they would've been in their conventional use)[10]. [0`auth`][0`part`] LFs retain their weak meaning "individual" (instead of being strengthened to "not-participant" as they would've been in their conventional use).

① Weak LF	② Stronger alternative(s)	③ Unstrengthened result
Participant singular $\lambda x : x \in \{i, u\} \,.\, x$ D ω [+`atom`] π [0`auth`][+`part`] ϕ	First singular $\lambda x : x \in \{i\} \,.\, x$ D ω [+`atom`] π [+`auth`][+`part`] ϕ	Participant singular $\lambda x : x \in \{i, u\} \,.\, x$ D ω [+`atom`] π [0`auth`][+`part`] ϕ "participant-but-not-author"
Individual singular $\lambda x : x \in \{i, u, o\} \,.\, x$ D ω [+`atom`] π [0`auth`][0`part`] ϕ	Participant singular $\lambda x : x \in \{i, u\} \,.\, x$ D ω [+`atom`] π [0`auth`][+`part`] ϕ	Individual singular $\lambda x : x \in \{i, u, o\} \,.\, x$ D ω [+`atom`] π [0`auth`][0`part`] ϕ "not-participant"

The semantic default in the category of number, which §3.2.3 established as the plural, is similarly derived. The same setup applies: in the contexts in which semantic defaults appear, there is nothing assumed about grammatical number in the conversation, and speakers are

[10]This makes the wonky prediction that this is participant-neutral, a weaker form of neutrality than being person-neutral. I do not know of any phenomena that are participant-neutral and not also person-neutral, but this is a prediction of the system that I make clear here.

unable to be more informative about this dimension of meaning. Recall our lottery scenario from (41), where the speaker is unable to assume a specific cardinality of winners. Only the number-neutral value is suitable. Since plural surfaces, it must be the number-neutral value. (If the number-specific value, singular, were used, that would be a breach of Quality.)

(41) *Context: A title of Mr. America is given to any contestant with scores above the 95th percentile. Nobody knows how many winners there will eventually be.*

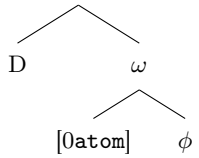
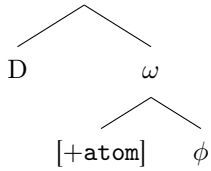
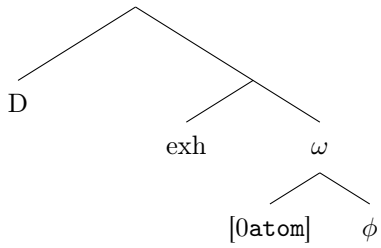
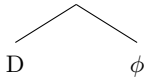
a. #The winner will be so lucky. It could be **me**.

b. The winner will be so lucky. It could be **us**.

Hence, semantic defaults for the category of number are defined by the absence of `exh` at ω. [0`atomic`] structures retain their number-neutral meaning (instead of being strengthened to "not-singular" as they would've been in their conventional use). This is illustrated for the semantic default of first person plural in (51b) below, but the same absence of strengthening at ω derives the plurals of second and third person.

① Weak LF	② Stronger alternative(s)	③ Unstrengthened result
Author number-neutral $\lambda x : x \in \{io*, iuo*\} \,.\, x$ D ω [0`atom`] π [+`auth`][+`part`] ϕ	Author singular $\lambda x : x \in \{i\} \,.\, x$ D ω [+`atom`] π [+`auth`][+`part`] ϕ	Author number-neutral $\lambda x : x \in \{io+, iuo+\} \,.\, x$ D ω [0`atom`] π [+`auth`][+`part`] ϕ "not-singular"

Okay. There are two more semantic defaults to derive: the inclusive and the impersonal. Recall these structures are posited to be structurally impoverished. These are semantic defaults because, by lacking the requisite functional heads, strengthening is *always* inoperative for number and/or person. On one hand, the impersonal inclusive may be strengthened for number when `exh` applies at ω, but may never be strengthened for person. It is partially unstrengthenable. On the other hand, the impersonal can never be strengthened for number or person, as its lack of functional projections means that there are no possible insertion sites for `exh`. It is completely unstrengthenable. These forms are shown below.

① Weak LF	② Stronger alternative	③ Strengthened LF
Individuals $\lambda x : x \in \{io*, iuo*, uo*, oo*\} \ . \ x$  	Singular person-neutral $\lambda x : x \in \{i, u, o\} \ . \ x$  	Plural person-neutral $\lambda x : x \in \{io+, iuo+, uo+, oo+\} \ . \ x$  "not-singular"
Individuals $\lambda x : x \in \{io*, iuo*, uo*, oo*\} \ . \ x$  	(none)	(none)

To summarize, semantic defaults are defined by an absence of strengthening via the absence of `exh` at ω and/or π. It is crucially different from their conventional use, which is defined by the presence of strengthening. In a privative feature system, zero-specified structures are born weaker than their positively-specified counterparts. Via competition, zero-specified structures may be strengthened, but *sans* competition, they retain their weak meanings.

What we know and love as "third person" conventionally is the semantic default for person, as it is born zero-specified as [0`auth`][0`part`]. Similarly, what we know and love as "plural" conventionally is the semantic default for number, as it is born zero-specified as [`atomic`]. Meanwhile, the defaulthood of the inclusive and impersonals is tied to the same combination of factors—zero-specification (by dint of lacking the requisite functional heads) and the absence of strengthening—so that they are semantic defaults for both person and number[11]. In contexts where there are no person and/or number assumptions being made (because of epistemic uncertainty, arbitrarity, or mixed quantification), only the number-neutral or person-neutral form may emerge.

Note that this account uses a combination of local strengthening (via the exhaustification operator `exh` (Chierchia, Fox & Spector 2021) and *Maximize Presupposition!*. This is because there are advantages for choosing this combined approach over approaches with only `exh` or only MP!.

A purely-`exh` approach will lack a mechanism for preventing alternatives which under-presuppose. In our combined approach, this is what *Maximize Presupposition!* fixes for us: amongst the available alternatives (if any), MP! chooses the one which is compatible with the context and is strongest presuppositionally.

In a purely-MP! approach, we would obtain conventional meanings in a slightly different

[11]Note that the weakest presupposition available in this system, $\lambda x : x \in \{io*, iuo*, uo*, oo*\} \ . \ x$, is one that makes no assumption being made about person *or* number, and is shared by third person plural and the impersonal. This explains why these ϕ-featural values are both attested in contexts of arbitrarity, as was shown in §3.3.

way. Consider the conventional meaning of third person, "not-participant". There would be competition between a sentence containing a first/second pronoun (which contributes the stronger presupposition that its referent be a participant) and a sentence containing a third person pronoun (which contributes no presupposition at all). If a speaker uses the latter sentence, this signals that the stronger presupposition does not hold, and that the negation of the stronger presupposition holds instead (the *anti-presupposition* of "not-participant"). This sentence-level competition is still perfectly capable of accounting for semantic defaulthood: one might posit that *Maximize Presupposition!* is simply inoperative in those special contexts allowing semantic defaults to arise, simply because there are no assumptions to be made about person and/or number. However, the purely-MP approach predicts that agreement is completely monolithic within the same sentence. This does not make the right predictions regarding the phenomenon of mixed agreement, which will be discussed later in §5.1.6.

On this note, we will now derive polite usages of pronouns. We will see that the analysis of polite usages almost perfectly resembles that which derived default usages, in that strengthening is similarly inoperative—but for different reasons. Specifically, strengthening is blocked by another pragmatic maxim, the Taboo of Presumptuousness.

4.4 Polite usages of pronouns: Strengthening has also defaulted

4.4.1 The gist: Redressive efforts in the face of Taboo

Our task now is to show that this system derives all and only the attested recruitment asymmetries for politeness established in Chapter 2. We will show that the relationship between proxy and proxand is similarly asymmetrical. Proxies are always more underspecified than their proxand, making them well-suited for expressing negative politeness.

Justifying the hullabaloo over semantic defaults in §3.2, I propose that polite proxies and semantic defaults are presuppositionally identical, being *person-neutral, number-neutral*, or *both*. Furthermore, both arise from contexts where strengthening is inoperative. For semantic defaults, strengthening was inoperative because the neutral form is already the most informative form; for proxies of politeness, strengthening is also inoperative because speakers are mandated to *not* be informative due to a Taboo on Presumptuousness, spelt out below. This is the crucial formal factor underlying their convergence, explaining why asymmetries in recruitment precisely align with asymmetries in semantic defaulthood.

It is worth emphasizing again that the notion relevant to the Taboo is that of semantic defaulthood, as neither syntactic nor morphological defaulthood bears on the range of meanings an element may have. What matters for the pragmatic maxim being proposed is presuppositional strength, not morphological complexity (as honorifics are form-identical to the features they recruit), or syntactic exceptionality (as honorific meaning does not appear or disappear with the type of construction used; neither does honorificity trigger certain syntactic operations).

The mechanism that enforces the use of underspecified forms is a pragmatic maxim termed the Taboo on Presumptuousness, repeated from (3). It is the antithesis of *Maximize Presupposition!*, repeated from (51). For the Taboo, presuppositionally strong forms are the worst choices possible; for *Maximize Presupposition!*, presuppositionally strong forms are the best choices possible. The Taboo is split into sub-maxims specialized for person and number in order to capture the parametric variation displayed by different languages in their choice of proxy. For instance, for the proxand of polite singular address, languages may choose a number proxy, a person proxy, or both. Importantly, languages do not *all* choose the most underspecified forms as the proxy; if that were the case, the only proxies that would be attested for all proxands would be 3PL and the impersonal, because these are both person-neutral and number-neutral. This prediction

is clearly false in light of the rich typological variation presented in Chapter 2. Hence, a one-size-fits-all Taboo would under-generate. (Later, in Chapter 4, the final Optimality-Theoretic analysis will generate a factorial typology which precisely mirrors the typology of polite pronouns from Chapter 2. There, the sub-Taboos will correspond to Markedness constraints.)

(3) Taboo on Presumptuousness (Taboo): Don't make assumptions about one's interlocutor in the conversation.

a. Don't make assumptions about the number of interlocutors (ω-Taboo).
b. Don't make assumptions about who the interlocutors are (π-Taboo).

(51) *Maximize Presupposition!* (MP!):
Choose the strongest presupposition compatible with what is assumed in the conversation.

Definitionally, then, polite contexts are contexts where the Taboo on Presumptuousness is operative. When the Taboo is operative, strengthening is unable to apply. Since the Taboo enforces that no assumptions be made about one's interlocutor, there is nothing assumed. As such, MP! is unable to enforce the use of a presuppositionally strong form over a weak one.

This implementation proposes that the degree to which the Maxim of Quantity is observed across (social) contexts is variable, as Quantity stands in tension with the Taboo. (Here, both `exh` and MP! are at the service of the more general Quantity.) Such tension was reflected in a debate about the universality of Grice's maxims, as highlighted by von Fintel & Matthewson (2008), where Keenan (1974: 258) claimed that speakers of Malagasy regularly violate the Quantity maxim:

> They regularly provide less information than is required by their conversational partner, even though they have access to the necessary information. If A asks B 'Where is your mother?' and B responds 'She is either in the house or at the market', B's utterance is not usually taken to imply that B is unable to provide more specific information needed by the hearer. The implicature is not made, because the expectation that speakers will satisfy informational needs is not a basic norm.

A reason that Keenan offers for this regular violation of Quantity is that individuals "do not want to be responsible for the information communicated." Additionally, Green (1990: 419) points out that the notion of being as informative as is required may be culturally variable. Being responsible for certain states of affairs may be an undesirable social outcome, so that interlocutors would rather violate Quantity than to assume responsibility for certain information.

Importantly, the recruiting mechanism regulating the expression of politeness simply piggybacks on the system responsible for semantic defaults; no machinery specific to politeness is required. The Taboo on Presumptuousness is indeed an innovation, but note that it is simply the linguistic reflex of a negative politeness strategy. The Taboo's preference for underspecification allows us to formally capture the intuition that avoidance is a key component of respect.

Note that (3) is a *pragmatic* maxim about speakers, stating what speakers may contribute to the conversation. An alternative *semantic* formulation may look like (52), where the maxim is about grammars instead. It penalizes the grammar for selecting strong linguistic forms (strengthened and/or positively-specified features in this system).

(52) Taboo on Strong Forms (ToSF): Don't use presuppositionally strong forms.

a. Don't use structures containing [+`atomic`] (ω-ToSF).
b. Don't use structures containing [+`auth`] or [+`part`] (π-ToSF).

In terms of predicting how recruitment is constrained, the two Taboos are equivalent: underspecified forms will be the forms recruited for politeness. However, I will favor the pragmatic formulation over the semantic formulation for the following reasons.

First, the pragmatic formulation is more parsimonious, and allows semantic defaults and proxies of politeness to arise from the same pragmatic landscape where *strengthening is inoperative*. Semantic defaults emerge when in environments (such as epistemic uncertainty, mixed quantification, arbitrarity) where strengthening cannot apply, as there is no maximal feature value being supplied by the context. Similarly, proxies for politeness emerge in environments where strengthening must not apply for negative politeness reasons. Hence, the two asymmetries converge because semantic defaults are definitionally meaning-poor; precisely these forms are recruited to approximate politeness *because* using a meaning-poor form *is* a way of being indirect. Such forms represent the closest approximations to a putative dedicated feature like [`indirect`] (or [`0direct`] that a grammar would possess in its inventory. Lest the reader think they are walking into a just-so equivalence between default and polite usages, the analysis will also be able to derive unattested proxies for politeness (because they are not semantic defaults) and cross-linguistic variation in the attested proxies for politeness (by selectively invoking the proposed pragmatic maxims).

Second, the pragmatic formulation has a better grip on deriving register-based variation in the same language; for instance, English *to eat* has many socially conditioned variants: formal *to partake, to dine* or informal *to tuck into, to nosh on*. Recall from §2 that some languages can display such variation in many more grammatical domains to create distinct speech levels, such as Javanese (4) which features different forms for pronouns and certain predicates across several speech styles: basic (*ngoko*), polite (*krama*), and super-polite (*krama inggil*). Since the application of the Taboo of Presumptuousness is dependent on the context of utterance, the assumption is that, for any context requiring politeness, it is operative and enforces the use of underspecified forms. Meanwhile, the application of the semantic Taboo on Strong Forms is dependent on the set of alternatives available language-wide, and assumes that they stay constant across all contexts of utterance, polite or non-polite. Contexts in which neutral forms are available would have to be separately stipulated, so that in *e.g.* Javanese, the Taboo on Strong Forms is operative only in the high *krama* levels, and not anywhere else. We would wrongly predict that speakers are unable to switch between speech levels in the same discourse depending on who their interlocutors are. Hence, the pragmatic formulation affords more flexibility to speakers, who may or may not decide to operationalize the Taboo depending on the speech situation.

The third reason for favoring the pragmatic formulation is that it explains why deniability—enabling the interlocutor to refuse—plays a crucial role in negative politeness. We repeat the relevant quote from Goffman (1967: 30) from §3.1.2 below:

> The rule regarding this unofficial kind of communication is that the sender ought not to act as if he had officially conveyed the message he has hinted at, while the recipients have the right and the obligation to act as if they have not officially received the message contained in the hint. Hinted communication, then, is deniable communication; it need not be faced up to.

By dint of being situated in a context in which speakers must not be informative about number/person, any proposition put forth about number/person *is deniable*.

For these reasons, we favor the pragmatic formulation of the Taboo in (3). We now examine the machinations of polite contexts, where (3) is dominant and forces the grammar to favor some forms over others. It is important to clarify that this "favoring" occurs at the level of DP. There

is no manipulation upstream to the valuation of the index by the assignment function: a polite utterance intended for some addressee will still actually refer to that addressee.

4.4.2 Weak forms as the output of Taboo

A theory of polite proxies must explain the recruitment asymmetries that we observed in Chapter 2, deriving all and only the attested recruitments for politeness, whether it be polite address, polite reference, or polite self-reference. We will consider each proxand in turn, as they are all different in terms of what is assumed in the conversation and what counts as most informative.

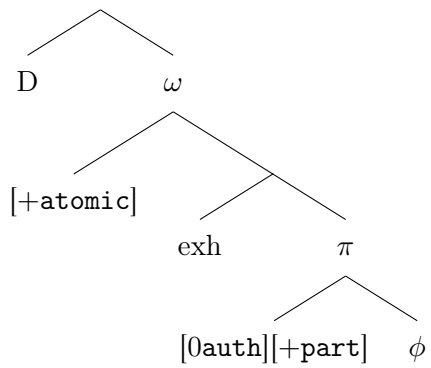
Our analysis of polite meaning will consist of two pragmatic layers. First, there is a layer of pragmatic reasoning in which the MP!-best form is arrived at, as in any conventional context. However, in a polite context, this MP!-best form is deemed offensive as it is only felicitous if the speaker makes assumptions about certain aspects of the context, violating the Taboo of Presumptuousness. Hence, there is a second, redressive layer of pragmatic reasoning in which the Taboo-best forms are favored over the MP!-best form. This redressive process consists of choosing forms which are presuppositionally weaker, which would normally result in violating MP!. Since polite contexts are definitionally those contexts in which the Taboo is operative, uninhibited by considerations of MP!, such violations are tolerated.

To illustrate this analysis, we start with the proxand of singular polite address.

4.4.2.1 Deriving polite address

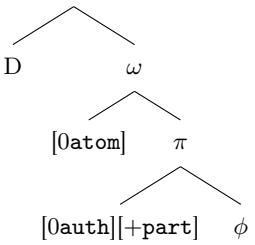
Consider the most frequent proxand, singular polite address, where what is assumed in the conversation is singular address. What goes wrong when the MP!-best form is used? This form has the structure in (53), and is offensive in two ways: [+atomic] is informative about number (singular), and the exhaustification at π is informative about person (participant-but-not-author).

(53) Offending structure when proxand = singular polite address

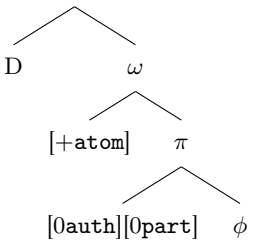


Hence, (53) represents the worst choice for the Taboo, so it is blocked from appearing as is in a polite context. Facing this dispreference, more Taboo-compliant alternatives—weaker alternatives—are called to the fore. These consist of structures where the number feature is zero-specified, (54a), where the person feature is zero-specified (54b), where both number and person features are zero-specified (54c), where the π head is simply not projected (54d), and where neither π nor ω are projected (54e). Additionally, there is no strengthening present throughout. Left unstrengthened, their neutral meanings are indicated below.

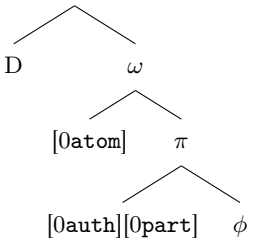
(54) Taboo-compliant alternatives to (53)

a. 

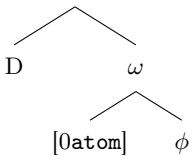
participant, number-neutral

b. 

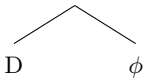
person-neutral, singular

c. 

person-neutral, number-neutral

d. 

person-neutral, number-neutral

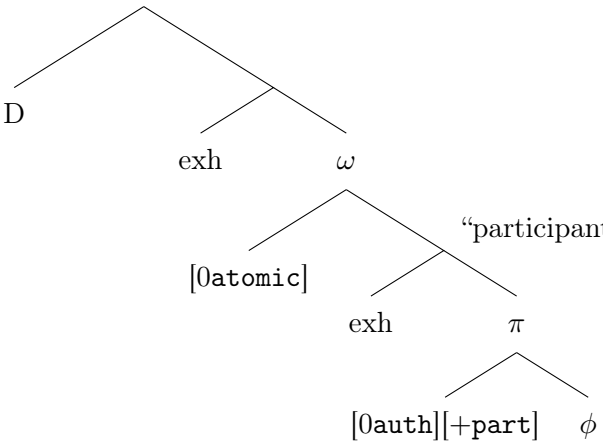
e. 

person-neutral, number-neutral

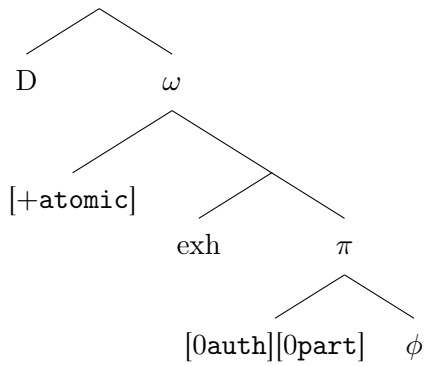
Crucially, their neutrality (on various dimensions) is what enables them to be proxies of polite address, as they are Taboo-compliant and well-suited for expressing negative politeness.

The weak alternatives in (54) correspond to the strengthened forms when `exh` is operative; when they are used conventionally. This is listed in (55).

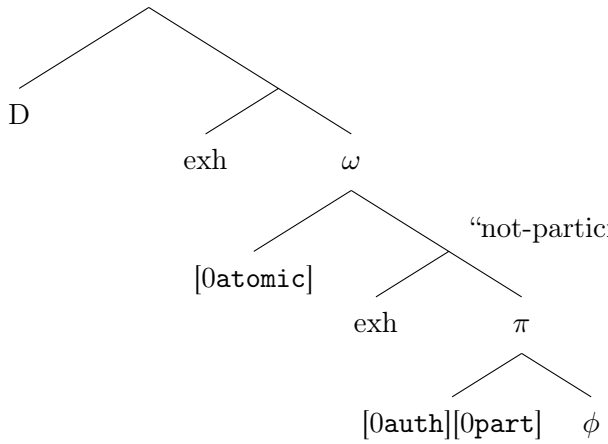
(55) Conventional counterparts of (54)

a. 

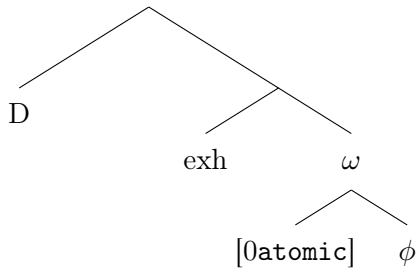
"participant-but-not-author; not-singular" (2PL)

b. 

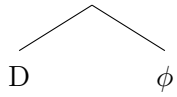
singular, "not-participant" (3SG)

c. 

"not-participant; not-singular" (3PL)

d. 

person-neutral, "not-singular" (our 1INCL.PL)

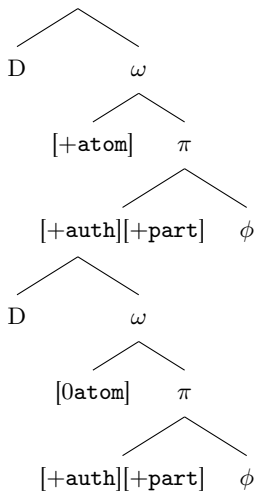
e. 

person-neutral, number-neutral (IMPERS)

Together, these unstrengthened (or unstrengthenable) alternatives comprise the set of attested proxies for the proxand of singular polite address, correspond precisely to those ϕ-featural values conventionally known as the second person plural, third person singular, third person plural, inclusive, and impersonal.

Conversely, stronger alternatives are *not* Taboo-compliant. Structures where the π head hosts a veritable swarm of restrictor functions—namely, [+auth][+part]—are *not* attested proxies for singular polite address. These are the structures in (56), which presuppose that their referent is an author; recall that the combination of [+auth] and [+part] necessarily restricts the ϕ-lattice to first persons. (Of course, the originally offensive structure in (53) is not a Taboo-compliant form either, but is not listed below.)

(56) Non-Taboo-compliant alternatives

a. 

author, singular

b.

author, number-neutral

These stronger alternatives, on the other hand, correspond precisely to those ϕ-featural values conventionally known as the first singular and the first plural, which are unattested proxies for this proxand. Hence, a polite grammar rules out alternatives which are too presuppositionally strong[12].

Crucially, the cut in the ϕ-featural space that (non-)strengthening makes mirrors exactly the cut made by the recruiting mechanism, as was established in Table 2.6, repeated below. We have now derived both sides of the recruitment asymmetry.

Recruitment asymmetries: Singular polite address								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	125	27	12	7	29
Language	–	–	–	French	Fehan Tetun	Italian	German	Ainu
				vous	*ita*	*Lei*	*Sie*	*an*

Having witnessed conventional, default, and polite ways of doing singular address, we will now illustrate the relevant derivations for these varieties of singular address in turn. In the following contexts, we entertain three male characters: Nigel, Caraway, and Gustavo, who all speak French and are fully aware of the *tu/vous* opposition.

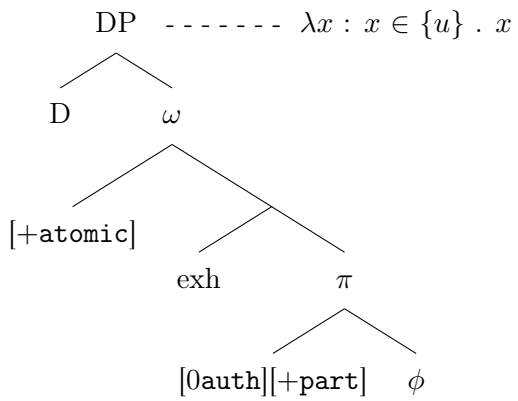
Throughout, the context furnishes the following assignment g as in (57).

(57) $[\![g]\!]^c = 3 \rightarrow$ Nigel; $4 \rightarrow$ Gustavo, $7 \rightarrow$ Caraway

Furthermore, we assume French makes available three LFs for second person, corresponding to varying presuppositions and spellouts, given in (59). (58a) presupposes that its referent is both singular and an addressee; (58b) presupposes that its referent is both plural and are addressees; (58c) presupposes that its referent is an addressee but does not presuppose anything about cardinality.

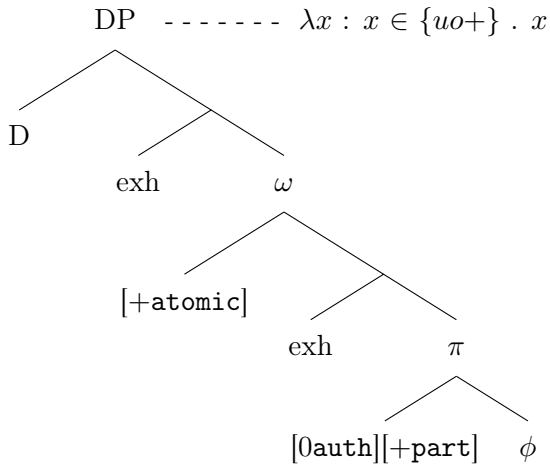
[12]Potentially, a polite grammar may also be defined as a sensible grammar, ruling out alternatives carrying presuppositions which clash with the pronoun's index. In this case, the pronoun's index refers to the addressee, clashing with the alternatives in (56) which presuppose the referent to be an author. In this conception, unattested proxies are unattested not because they are too strong, but rather because they are too insensible. However, I will go with the former too-strong conception because it is more compatible with the impending OT analysis in the next chapter: when Markedness constraints, proposed as OT counterparts of the Taboo, are ranked high enough, unattested proxies are ruled out. Implementing the too-insensible conception is possible as well, with a putative NoClash constraint responsible for unattested proxies instead of Markedness, but is not pursued here.

(58) a.



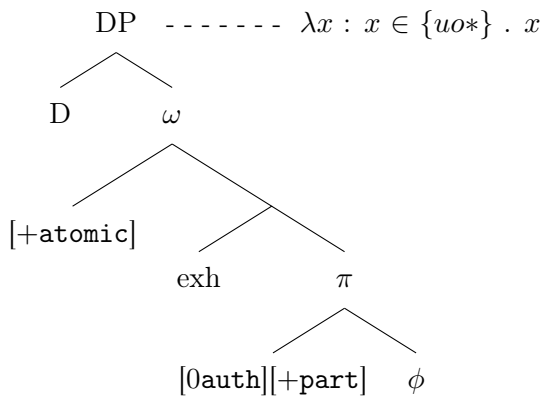
$\Leftrightarrow$ *tu*

b.



$\Leftrightarrow$ *vous*

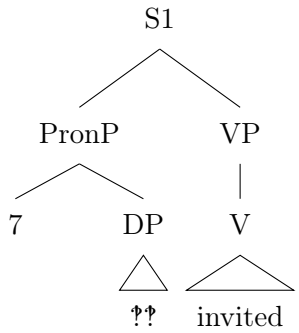
c.



$\Leftrightarrow$ *vous*

In the first scenario, Nigel would like to invite his childhood friend Caraway to dinner. They have known each other since kindergarten, and are on extremely informal terms. We assume that Nigel's invitation occurs in the context c_1 (59a), and has the syntactic structure consisting of a PronounP and a VP as in (59b). The contents of DP are indicated for now with ⮝⮝, as the selection of a particular DP out of the three illustrated depends on the pragmatic maxims operative in the context, and will be discussed soon enough.

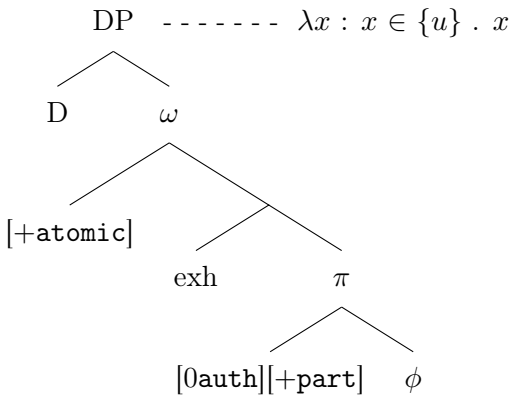
(59) a. c_1: referent of the pronoun is a singular addressee.
b.



The truth conditions of S1 are straightforward and are spelt out in (60). Since the pronoun refers to Caraway, the assignment function points to the index 7, according to (57).

(60) $[\![\text{S1}]\!]^{w,g} = [\![\text{VP}]\!]^{w,g}\ ([\![\text{PronP}]\!]^{w,g})$
$= [\![\text{be-invited}]\!]^{w,g}\ (g(7))$
$= (\lambda x\ .\ \text{be-invited}_w(x))\ (g(7))$
$= \text{be-invited}_w(g(7))$

The three available DPs in (58) present MP! with a choice to make. Amongst them, MP! prefers the one that has the strongest satisfied presupposition. (58a), repeated below, carries a presupposition that matches c_1 the most closely, that the referent is an addressee and is singular. Since (58a) can be used, MP! states that it must be used. (For MP!, (58a) wins out over (58b) and (58c), as these DP competitors carry less-strong presuppositions with regards to number. Both presuppose that the referent is an addressee, but (58b) presupposes its referent is plural and (58c) makes no presupposition about cardinality.)

(61)


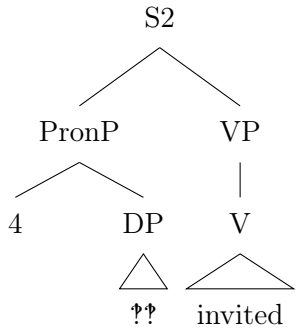
⇔ *tu*

Furnished with the appropriate verbal agreement, Nigel's invitation to his friend Caraway might resemble something like (62).

(62) *Tu es invité.*
2SG COP.2SG invite.PAST.PART.MASC.SG
'You're invited (to dinner tomorrow).'

Now let us consider a second scenario, where Nigel would like to invite his father-in-law Gustavo to dinner for the first time. Nigel and Gustavo have not been acquaintances for a long time, and owing to their social relation, they are on extremely formal terms. We assume that Nigel's invitation occurs in the context c_2 (63a), and has the same syntactic structure as in (63b).

(63) a. c_2: referent of the pronoun is a singular addressee *to whom deference must be paid.*
b.



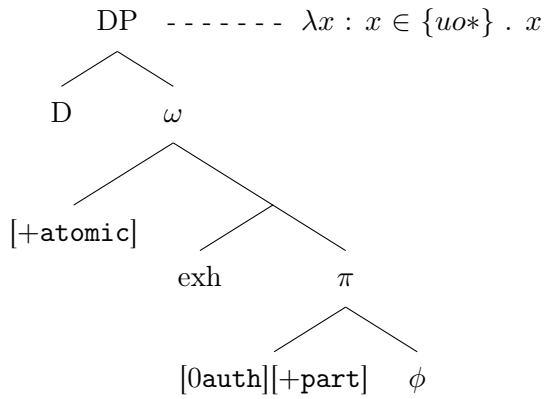
The truth conditions of S2 are spelt out in (64). Since the pronoun refers to Gustavo, the assignment function points to the index 4, according to (57).

(64) $[\![\text{S2}]\!]^{w,g} = [\![\text{VP}]\!]^{w,g}\ ([\![\text{PronP}]\!]^{w,g})$
$= [\![\text{be-invited}]\!]^{w,g}\ (g(4))$
$= (\lambda x\ .\ \text{be-invited}_w(x))\ (g(4))$
$= \text{be-invited}_w(g(4))$

In this *polite* context, which of the three DPs is appropriate now? In order to satisfy *Maximize Presupposition!*, the appropriate DP must—if feasible—reflect aspects of c_2, namely, that the referent is singular and an addressee to whom deference must be paid. Out of the three DPs in (58), (58a) (which won in c_1 above) would be favored by MP!.

However, (58a) is *not* appropriate here, as it is only felicitously used in contexts where there *are* assumptions made about the cardinality of the interlocutor. As a ω-Taboo is operative in this polite context c_2 ("Don't make assumptions about the number of interlocutors"), the number-neutral DP (58c), repeated below, wins instead.

(65)



⇔ *vous*

Note that this Taboo-favored LF is also compatible with a context where Nigel is unable to make any assumptions about cardinality: not because of being in a polite context and having to be Taboo-complaint, but rather due to e.g. epistemic uncertainty. This explains the convergence between the shapes of polite proxies and semantic defaults.

Furnished with the appropriate verbal agreement, Nigel's invitation to his father-in-law Gustavo might resemble something like (66).

(66) *Vous êtes invités.*
2PL COP.2PL invite.PAST.PART.MASC.PL

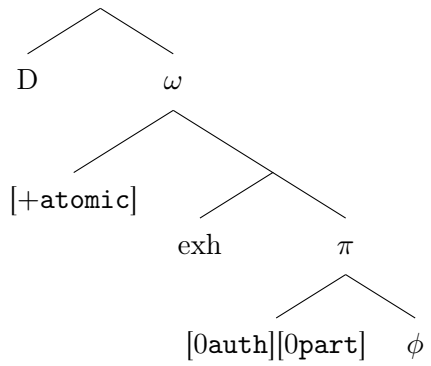
'You're invited (to dinner tomorrow).'

Above, we have pretended that the derivation of polite address is independent of that of non-polite address, in the sense that the DP chosen for S2 does not hinge upon the DP chosen in S1. However, it is just that—a pretense! We said at the beginning of this section that our analysis of polite meaning will consist of two pragmatic layers, and this is precisely what is implemented in Chapter 5, where polite meaning is an additional, redressive layer atop conventional meaning modulated by Optimality-Theoretic considerations. In the first conventional layer, the MP!-best form wins. This form serves as the input which must be "redacted" in some way when it serves time in the second polite layer, which produces forms which are weaker in meaning in various ways.

4.4.2.2 Deriving polite reference

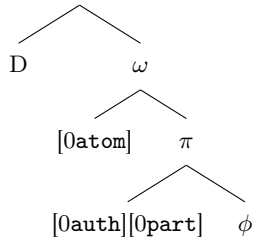
Let us now consider another proxand, singular polite reference. Here, what is assumed in the conversation is singular reference. What goes wrong when the maximally MP!-compliant pronoun is used? This has the structure in (67), and is offensive in two ways: [+`atomic`] contributes information about number (singular), and the exhaustification at π contributes information about person by winnowing down the ϕ-space from the domain of all individuals to the not-participant subregion.

(67) Offending structure when proxand = singular polite reference

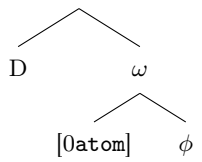


Hence, (67) is the least Taboo-compliant form. The redressive strategy is as before: seek alternatives which feature weaker meanings. Suitable alternatives are structures where the only positively-specified feature [+`atomic`] is now zero-specified (68a), where the π head is simply not projected (68b), and where neither π nor ω are projected (68c). Additionally, there is no strengthening present throughout, so that they have the neutral meanings indicated below.

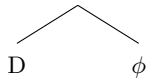
(68) Taboo-compliant alternatives to (67)

a. 

person-neutral, number-neutral

b. 

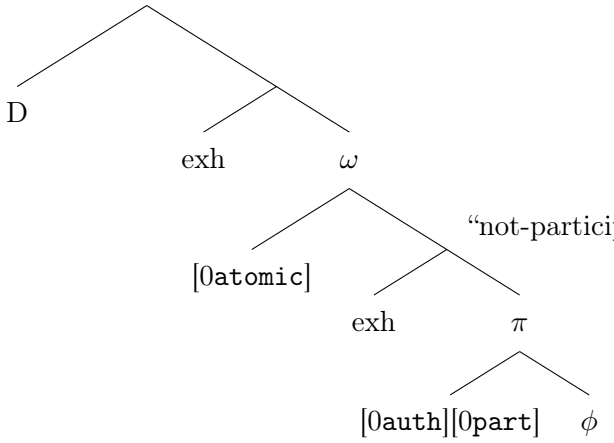
person-neutral, number-neutral

c. 

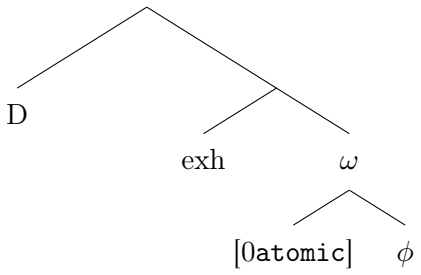
person-neutral, number-neutral

Again, in the absence of strengthening, the structures in (68) are neutral with respect to person *and* number, making them Taboo-compliant. In the presence of strengthening, the weak LFs from (68) correspond to the following given in (69): the third person plural, inclusive, and impersonal, respectively.

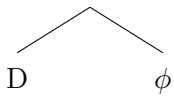
(69) Conventional counterparts of (68)

a. 

"not-participant; not-singular" (3PL)

b. 

person-neutral, "not-singular" (our 1INCL.PL)

c. 

person-neutral, number-neutral (IMPERS)

Together, these unstrengthened (or unstrengthenable) alternatives comprise the set of attested proxies for the proxand of singular polite reference, correspond to those ϕ-featural values conventionally known as the third person plural, inclusive, and impersonal.

As before, stronger alternatives to the originally offensive structure are *not* Taboo-compliant. (Of course, the originally offensive structure in (67) is not a Taboo-compliant form either.) These are LFs positively specified for person features, either as [+`auth`][+`part`] (restricting the ϕ-lattice to authors) or as [0`auth`][+`part`] (restricting the ϕ-lattice to participants; when `exh` applies, further restricting the ϕ-lattice to addressees). Crucially, these stronger alternatives correspond precisely to those ϕ-featural values conventionally known as the first and second persons, which are indeed unattested proxies for this proxand.

It is worth checking whether our analysis mirrors the cut made by the recruiting mechanism for this proxand as established in Table 2.7, repeated below.

Recruitment asymmetries: Singular polite reference								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	0	0	0	28	8
Language	–	–	–	–	–	–	Wolaytta	W. Apache
							ʔetaa	*ko-*

Almost: We immediately notice that the inclusive is wrongly predicted to be attested at this point; however, it will be eliminated by the impending Optimality Theory analysis in the next chapter. The rest of the recruitment asymmetry is correctly predicted: third person plurals and impersonals are attested proxies, while first and second persons are unattested proxies. For now, we will move on to the last remaining proxand, that of polite/humble self-reference.

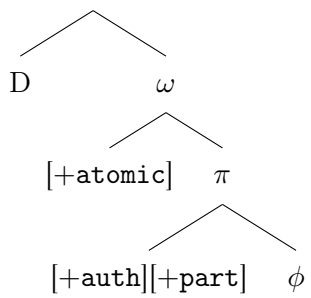
4.4.2.3 Deriving polite/humble self-reference

The proxand of polite self-reference is vanishingly rare, with only ≈10 instances in the typology. The recruitment asymmetry it exhibits is repeated from Table 2.8 below.

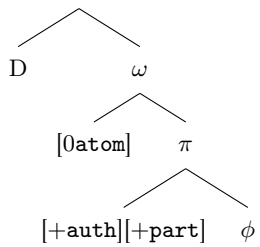
Recruitment asymmetries: Singular polite self-reference								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	6	0	0	4	0	0	1
Language	–	Persian	–	–	Guarijio	–	–	E. Kayah
		mā			*remé*			*ʔū*

What is assumed in the conversation in this case is singular self-reference. The most MP!-compliant pronoun in this case, 1SG, involves the maximum number of positive specifications possible in the proposed feature system, where all three features are positively-specified as in (70).

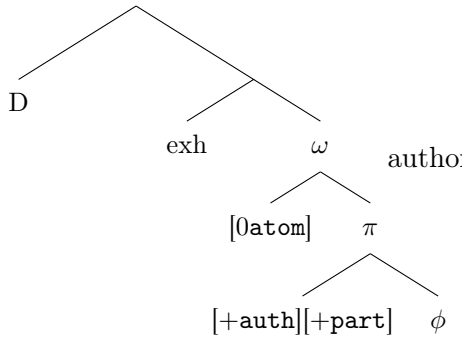
(70) Offending structure when proxand = singular polite self-reference



Within its set of attested proxies (1PL, inclusive, and impersonal), 1PL is its most robustly attested proxy, and is easily derived. A minimally different Taboo-compliant alternative would be (71), which is just a bit more meaning-poor by featuring zero-specification at ω, making it number-neutral. Its conventional counterpart (72) would feature strengthening at ω, so that its number-neutrality is strengthened to "not-singular". This explains the shape of the 1PL proxy.

(71) 

author, number-neutral

(72) 

author, "not-singular" (1PL)

What about its other proxies? Since (70) is the most positively-specified structure possible in the proposed system, its weaker alternatives comprise a veritable panoply: any plural (via [0atomic]), any non-author form (via [0auth]), and any non-participant form (via [0part]). So far, we predict that this proxand should have available the most numerous proxies available—but, sadly, contrary to fact. Perhaps they are just obscured by the rarity of the proxand, but I will argue in the next section that this is an effect of Faithfulness constraints: the recruiting mechanism tolerates deviations from the form closest to the proxand, but only to a certain extent. Hey, wait a minute—Faithfulness, you say? Isn't that Faithfulness from Optimality Theory? Yes, it is.

Lest the reader is getting weary, the Optimality Theoretic-analysis is able to illuminate *which* less-informative alternative, out of a set of several possible alternatives, emerges in a particular language, and also explain the shapes of multiple-tiered politeness systems. So far, we have been able to describe the constraints on possible proxies, but unable to explain why one proxy wins out over the others. After all, most languages typically pick one proxy for politeness instead of offering multiple options. Worse, we have completely ignored the asymmetries inherent in multiple-tiered systems, which §2.2 spent much time dwelling on.

The Optimality Theoretic-analysis will be able to describe the conditions under which one proxy wins out over the others in terms of the negative politeness strategy preferred by the language, and also distinguish proxies by how much negative politeness they can express. With these promises in mind, let us go together there now.

Chapter 5

The Final Analysis

So far, we have been invoking a "recruitment mechanism" but have not yet made its machinations transparent. As it stands, it predicts that *some* ϕ-featural values may be recruited, but does not predict *which* specific value may be recruited for *which* tier of politeness.

This chapter takes us into an Optimality-Theoretic implementation where concrete predictions are made about *which* neutral ϕ-featural value may be chosen for recruitment, and under *which* setting. The specifics of the recruitment mechanism are hence illuminated. This analysis for proxies of politeness has more predictive power than what we were pursuing in the previous chapter: §5.1 derives all and only the attested proxies for all proxands in the typology, and explains why politeness clines have the shape that they do. The resulting factorial typology is exactly parallel to the typology of politeness presented in Chapter 2. (If the previous chapter were left as is, we would predict only two kinds of typological scenarios: the *free-for-all* and the *zero-for-all*. In the *free-for-all* scenario, all neutral ϕ-featural values should be recruited to be in free variation as proxies within a single language. However, languages like French only make available one proxy for singular polite address (the character conventionally known as 2PL), and not the full panoply of cross-linguistically attested proxies (the characters conventionally known as 2PL, 3SG, 3PL, INCL, and the impersonal). In the *zero-for-all* scenario, languages should only recruit the LFs with the least restrictive meanings (the characters conventionally known as 3PL and the impersonal) for politeness, because these are the most Taboo-compliant. Neither scenario pertains, so there must be an additional component deriving the subtler nuances of typological variation.)

In light of this final analysis, §5.2 considers some surprising gaps in recruitment from other ϕ-featural categories. §5.3 compares the proposed analysis to existing analyses of politeness and honorification, its most notable difference being the lack of a feature dedicated to politeness like [HONORIFIC] or [+`formal`] situated in the core grammar.

5.1 Optimality Theory: Polite and non-polite grammars

Optimality Theory is a natural analytical choice here because it evaluates the suitability of several candidates relative to each other. Hence, within a pool of several candidates, constraint rankings can vary to favor one candidate over another. This will be used to explain the politeness clines established in §2.2. It also offers a powerful tool for eliminating cross-linguistically unattested candidates: harmonic bounding. If a candidate C_A can never arise under any constraint ranking because it incurs more violations than C_B, then C_A is said to be harmonically bound by C_B. We will utilize its predictive power, and show that harmonically bound candidates are always unattested proxies of politeness.

In this section, we will differentiate between polite grammars and non-polite grammars, situating the difference in constraint ranking. Polite grammars are those constraint rankings which give rise to attested proxies of politeness, and non-polite grammars are those constraint rankings which give rise to unattested proxies of politeness. In the previous chapter, we proposed the following pragmatic maxims in (3) and (51). We defined *polite contexts* as those contexts where the family of Taboo maxims are operative, thereby bleeding the conditions of application for MP!, such that MP! is inapplicable.

(3) Taboo on Presumptuousness (ToP): Don't make assumptions about one's interlocutor in the conversation.

  a. Don't make assumptions about the number of interlocutors (ω-Taboo).
  b. Don't make assumptions about who the interlocutors are (π-Taboo).

(51) *Maximize Presupposition!*:

Choose the strongest presupposition compatible with what is assumed in the conversation.

The OT counterparts of the Taboo in (3) are the following Markedness constraints in (1). (1a) penalizes every instance of [+auth] or [+part] (corresponding to the π-Taboo; in this framework, to make assumptions about person necessarily involves positively-specified person features). (1b) penalizes every instance of [+atomic] (corresponding to the ω-Taboo; in this framework, to make assumptions about number necessarily involves a positively-specified number features). These militate against positively-specified features, which results in stronger restrictions upstream. Since the Taboo on Presumptuousness is operative only in polite contexts, its corresponding Markedness constraints are specified to apply only in polite contexts, and do not assign violations in non-polite contexts.

(1) Markedness constraints

a. *Indexical (*Index): Penalize every instance of [+F] hosted under the person head π in polite contexts.

b. *Individuate (*Indiv): Penalize every instance of [+F] hosted under the number head ω in polite contexts.

On the other hand, the Faithfulness constraints in (2) are the OT counterparts of *Maximize Presupposition!* in (51). They serve to militate against deviations from the input in the output, keeping Markedness in check. (2a) militates against any change in feature specification between correspondents in the input and output. (2b-c) militate against any deletion of ω and/or π projections from the input. Since (positive) feature specification essentially narrows down the space of possible meanings, these Faithfulness constraints urge the output to be as informative about the input as possible.

(2) Faithfulness constraints

a. IDENT(F): Penalize every change in feature specification from the input.

b. MAX(π): Penalize every deletion of a π projection from the input.

c. MAX(ω): Penalize every deletion of a ω projection from the input.

Note that IDENT(F) is sensitive only to changes in features between correspondents, *i.e.* when [+F] in the input corresponds to [0F], or *vice versa*. IDENT(F) is impervious to the wholesale deletion of features, which occurs when functional heads hosting those features are deleted. Since deletion destroys the correspondence relation, IDENT(F) is vacuously satisfied[1].

Constraints are ranked relative to each other to produce different grammars, whether they be polite grammars or non-polite grammars. In general, the more a candidate satisfies Markedness constraints, the more polite the tier it is recruited for will be.

Right now, we will focus on the relative rankings in polite grammars. The core advantage of OT is that, given a delineated pool of distinct outputs, it allows us to understand *which* output is the most ideal under a given ranking. For instance, we have proposed that the impersonal and third person plural are both completely presuppositionless for person and number. So, why

[1]Since [0F] is an absence of a feature and [+F] is the presence of a feature, some might object that the correspondence of [+F] in the input to [0F] should not incur a true IDENT(F) violation. Perhaps a better constraint here would be MAX(F), which this correspondence would incur a true violation of. However, using MAX(F) runs us into another problem: it does not distinguish deleting a feature *vs.* deleting a functional projection hosting that feature. For a [+part] input, for example, it would not distinguish between a [0part] output (the number-neutral candidate) and an output that does not host π outright (the inclusive candidate), even though there is cross-linguistic variation attested there. Hence, we will use IDENT(F), and simply assume that the correspondence of [+F] in the input to [0F] *does* incur a true IDENT(F) violation.

do they not emerge always as the only proxy of politeness for every language? To the contrary, the proxy of the second person plural is the most frequent, even though it is presuppositionless only for number.

We will see that the interplay of Markedness with Faithfulness constraints will allow us to adapt gradient preferences for different candidates, allowing a powerfully precise grip on cross-linguistic variation. Since specific rankings generate specific outcomes, this affords a grammar the choice between recruiting plural, third person, inclusive, or impersonal. Abstracting over the vagaries, we can identify common polite grammars that result in specific recruits being chosen for politeness, *polite grammars*, which are schematized in (3). (3a) results in the plural being recruited: the language would rather delete positive specifications than to individuate. (3b) results in the third person: the language would rather delete positive specifications than to use an indexical. (3c) results in the third person plural: the language would not like to individuate or to use an indexical. (3d) results in the inclusive: the language would rather delete the π projection than to host positive specifications there. Lastly, (3e) results in the impersonal: the language deletes both ω and π projections at all costs, consequences be damned. (3) thus represents a set of tradeoffs between negative politeness (Markedness) and Gricean Quality (Faithfulness).

(3) Polite grammars

a. Proxy = plural when MAX(π), MAX(ω) > *Indiv > ID(F)
b. Proxy = 3SG when MAX(π), MAX(ω) > *Index > ID(F)
c. Proxy = 3PL when MAX(π), MAX(ω) > *Indiv, *Index > ID(F)
d. Proxy = inclusive when *Indiv, *Index > MAX(ω) > ID(F), MAX(π)
e. Proxy = impersonal when *Indiv, *Index > ID(F) > MAX(ω), MAX(π)

Note also that (3a-e) escalate in number of Faithfulness violations: (3e) incurs the most while (3a) incurs the fewest. This means that, (3e) is a grammar that generates the highest degree of politeness, while (3a) is a grammar that generates the lowest degree of politeness (without sinking into a non-polite grammar). Indeed, the gradients between polite grammars will help us explain the politeness clines established in §2.2.

There is a general question of why *Indiv > ID(F) in (3a) is a less polite grammar than *Index > ID(F) in (3b) is, as this seems to be in cahoots with the fact that number proxies are always recruited for lower tiers than person proxies. I argue that this arrangement is principled and trace the difference between number and person to a difference in indexicality. Grammatical persons are indexicals: the reference of a personal pronoun like *I* or *you* crucially refers to the context parameter. In contrast, numbers are not *indexicals*, a class of expressions whose reference changes with the context of utterance. The reference of a plural-marked noun like *cats* does not depend on any aspect of the context. Furthermore, in surveys of indexical shift (e.g. Deal 2020), participant pronouns frequently participate in this phenomenon, whereas number is never found to participate. I would like to suggest that indexicals are less suitable for indirectness than non-indexicals are (cf. Wang 2025). This is because there is a dedicated mechanism in the grammar—the context parameter—that fixes their reference. Meanwhile, there is no comparable grammatical mechanism that fixes the reference of an individual within a plurality, so that using a mismatch in person contributes more negative politeness than using a mismatch in number.

Another abstraction is represented by the relative ranking between ID(F) with MAX(ω) and MAX(π). In (3a-c), ID(F) is outranked by MAX(ω) and MAX(π); in (3d-e), the reverse pertains such that MAX(ω) and MAX(π) outrank ID(F). Why are (3d-e) more polite grammars than (3a-c)? I suggest that it is because choosing a structurally poor form is an even more polite move than choosing a denotationally poor form is: structurally poor forms lack functional heads altogether, precluding the possibility of strengthening. This is another way of obeying the Taboo,

which militates against using forms with more presuppositional strength. Since structurally poor forms are partially or completely unstrengthenable, they are highly Taboo-compliant.

All other rankings result in *non-polite grammars*, such as those in (4). A notable non-polite grammar is (4a), where all Faithfulness constraints outrank all Markedness constraints. This ranking will simply produce the most faithful candidate, e.g. 2SG for singular polite address. While this is certainly a possible outcome—there are plenty of languages which do not recruit any pronoun for singular polite address—it is not a polite outcome. There are several more non-polite grammars: such as those which would rather use an indexical rather than deleting feature specifications (4b), or use a singular rather than deleting feature specifications (4c). These are the unattested languages of kipuY (43) and hcnerF (42), respectively, languages which use first person and singular, to indicate politeness, which are indeed unattested.

(4) Non-polite grammars

a. MAX(π), MAX(ω), ID(F) > *Indiv, *Index
b. ID(F) > *Index > MAX(π), MAX(ω)
c. ID(F) > *Indiv > MAX(π), MAX(ω)

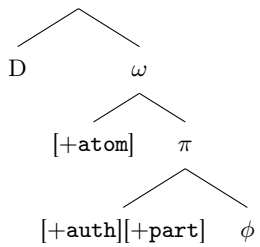
We will see that it is the *intercalated* outranking of Markedness over Faithfulness that derives the desired factorial typology. We retain the core insight of the previous chapter—that a polite context is one where speakers do not make assumptions, where presuppositionless forms are preferred over presupposition-rich ones. The degree of presuppositionless, however, is parameterized by OT, and explains typological variation.

Lastly, we come to the question of how a particular representation is anointed as the input (or Underlying Representation, UR). The official answer is that arriving at a polite outcome involves two layers of pragmatic reasoning. First, the form that is MP!-best (the form that is conventionally used in a non-polite context) is chosen. However, since the Taboo is the dominant pragmatic maxim in a polite context, a second layer of reasoning seeks alternatives to the MP!-best form according to various constraint rankings. Thus, a polite grammar is an additional, redressive layer atop the conventional grammar. In the following, inputs will be those structures which is the best alternative for *Maximize Presupposition!*, while outputs are those alternatives which are well-suited alternatives for the Taboo. This two-layered approach—with polite meaning as a "redacted" layer atop conventional meaning—was illustrated in §4.4.2.1 in the last chapter. In this chapter, we will go through the nuances of this redaction.

Now we are equipped to derive the proxies for polite address, reference, and self-reference. I will start with deriving proxies of polite self-reference, a conundrum we concluded with in Chapter 4. While it is vanishingly rare, the input, 1SG, is highly positively-specified and therefore highly unsuitable for politeness. Considering its less positively-specified alternatives will help us appreciate the effects of ID(F).

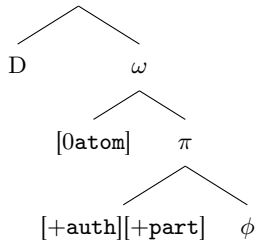
5.1.1 Deriving proxies of singular polite self-reference

While the data on this proxand is thin, we consider it first to see how the OT analysis works since the relevant 1SG input is the LF with the most restrictive meaning in the arsenal of LFs, and thus the biggest bane of the Taboo's existence. All three features are positively specified as [+`atomic`], [+`author`], and [+`participant`].

(5) 

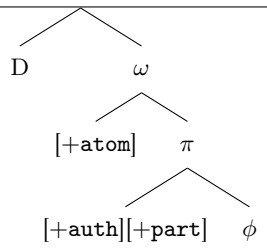
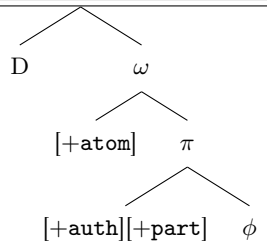
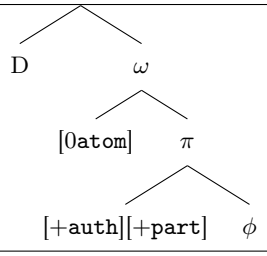
(⇒1SG)

Let us consider an alternative LF which is slightly weaker in meaning. (6) denotes a number-neutral first person by dint of being [0atomic], which would be strengthened to first person plural in its conventional use if *Maximize Presupposition!* were operative. As it is being used politely, LFs are left unstrengthened, so it is not strictly correct to talk about (6) as a "first person plural" but rather as a "first person number-neutral" form. For ease of exposition, however, the notation (⇒X) indicates its conventional name.

(6) 

(⇒1PL)

Let us evaluate (6) as an alternative to (5), involving a minimal change in number to [0atomic]. Throughout the tableaux in this chapter, the first candidate, listed in (a), will always be the faithful candidate. Below, (6) is listed as candidate (a) and (5) as candidate (b). (b) incurs a violation of ID(F) because [+atomic] in the input has been changed, resulting in [0atomic]. Meanwhile, (a) violates *Indiv by dint of hosting a positively-specified feature under ω, [+atomic]. When *Indiv outranks ID(F), (b) will win; otherwise, the faithful candidate (a) will win. (Both candidates violate *Index to the same degree, by dint of hosting two positively-specified features under π.) This is displayed in Table **I**.

		MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
I						
a	 (⇒1SG)	0	0	**	*!	0
$\boxed{\text{b}}$	 (⇒1PL)	0	0	**	0	*

Let us also evaluate an alternative to (5) that involves an additional change in person, resulting in not only [0`atomic`] but also [0`author`]. This candidate, (b), incurs two violations of ID(F) because [+`author`] and [+`atomic`] in the input has been changed to their zero-specified counterparts. This ensures that (b) features fewer violations of *Index and *Indiv than (a), which violates *Indiv once with its [+`atomic`] and *Index twice with its [+`auth`][+`part`]. When *Index *or* *Indiv outrank ID(F), (b) will win; otherwise, the faithful candidate (a) will win. This is displayed in Table **II**.

			MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
II	D, ω; [+`atom`], π; [+`auth`][+`part`], ϕ						
a	D, ω; [+`atom`], π; [+`auth`][+`part`], ϕ	(⇒1SG)	0	0	**!	*!	0
(b)	D, ω; [0`atom`], π; [0`auth`][+`part`], ϕ	(⇒2PL)	0	0	*	0	**

Lastly, let us evaluate an alternative which involves changing both positively-specified person features to zero-specified ones, resulting in [0`author`] and [0`participant`]. This candidate, (b), incurs two violations of ID(F) because [+`author`] and [+`participant`] in the input has been changed to their zero-specified counterparts. This ensures that (b) features no violations of *Index whatever, while (a) racks up two *Index violations. In this scenario, (b) will win when *Index *or* *Indiv outrank ID(F). Otherwise, the faithful candidate (a) will win. This is displayed in Table **III**.

III	D ω [+atom] π [+auth][+part] ϕ		MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
a	D ω [+atom] π [+auth][+part] ϕ	(⇒1SG)	0	0	**!	*!	0
$\boxed{\text{b}}$	D ω [+atom] π [0auth][0part] ϕ	(⇒3SG)	0	0	0	*	**

Now that we have illustrated the competition logic between individual candidates to the input, they are all compared relative to each other in the tableaux below. As always, (a) is the faithful candidate which is maximally positively-specified. (b) features a minimal change from [+atom] to [0atom]; (c) features a minimal change from [+auth] to [0auth]; (d) features more changes, from [+atom][+auth] to [0atom][0auth], retaining only one positively-specified feature [+part]; (e) retains only one positively-specified feature [+atom]; (f) is maximally zero-specified.

We first note two instances of *collective harmonically bounding*, where several candidates gang together to beat one candidate "on some of the possible constraint rankings while leaving none uncovered" (Samek-Lodovici & Prince 2002). Such a candidate is (c), which is collectively harmonically bound because it incurs one violation each of *Indiv, * ID(F), and *Index. All other candidates feature at least one non-violation of these constraints, beating (c) on this account, so that (c) can never arise. (d) is also collectively harmonically bound, by (b) and (f): (d) has more violations for ID(F) compared to (b), but also more violations for *Index compared to (f). Since (b) and (f) will always be more ideal than (d) in terms of ID(F) and *Index, (d) can never arise either.

The remaining candidates, (a, b, e, f) are predicted to be possible outcomes. We will go through the rankings deriving them in turn. Under the current ranking where ID(F) is lowest-ranked, (f) wins, as in Tableaux **IV**.

However, there are multiple rankings in which (f) loses. One way for (f) to lose is when *Indiv outranks ID(F) and *Index. In this case, (f) loses to (b). This is displayed in Tableaux **V**.

			MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
IV	D ω [+atom] π [+auth][+part] ϕ						
a	D ω [+atom] π [+auth][+part] ϕ	(⇒1SG)	0	0	**	*	0
b.	D ω [0atom] π [+auth][+part] ϕ	(⇒1PL)	0	0	**	0	*
c	D ω [+atom] π [0auth][+part] ϕ	(⇒2SG)	0	0	*	*	*
d	D ω [0atom] π [0auth][+part] ϕ	(⇒2PL)	0	0	*	0	**
e	D ω [+atom] π [0auth][0part] ϕ	(⇒3SG)	0	0	0	*	**
$\boxed{\text{f}}$	D ω [0atom] π [0auth][0part] ϕ	(⇒3PL)	0	0	0	0	***

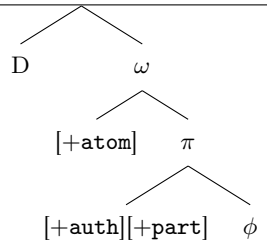
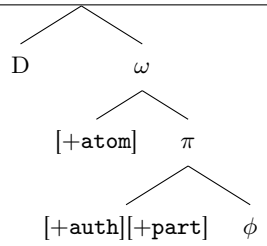
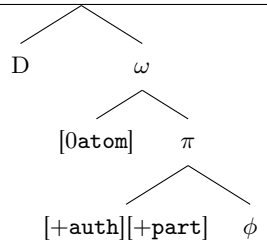
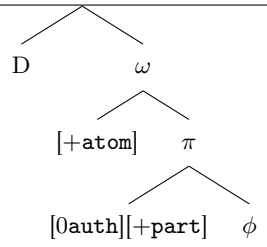
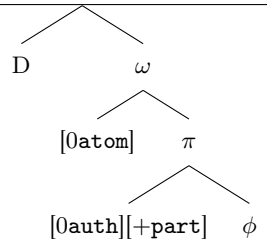
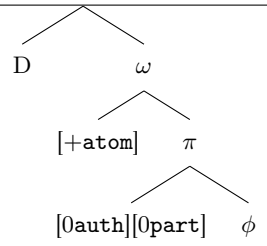
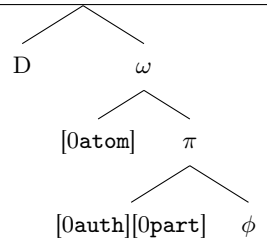
			Max(ω)	Max(π)	*Indiv	Id(F)	*Index
V	D ω [+atom] π [+auth][+part] ϕ						
a	D ω [+atom] π [+auth][+part] ϕ	(⇒1sg)	0	0	*!	0	**
$\boxed{\text{b}}$	D ω [0atom] π [+auth][+part] ϕ	(⇒1pl)	0	0	0	*	**
c	D ω [+atom] π [0auth][+part] ϕ	(⇒2sg)	0	0	*!	*	*
d	D ω [0atom] π [0auth][+part] ϕ	(⇒2pl)	0	0	0	**!	*
e	D ω [+atom] π [0auth][0part] ϕ	(⇒3sg)	0	0	*!	**	0
f	D ω [0atom] π [0auth][0part] ϕ	(⇒3pl)	0	0	0	***!	0

Another way for (f) to lose is for Id(F) to outrank both *Indiv and *Index. In this case, the faithful candidate (a) wins. Yet, this is *not* a polite grammar as we are reminded by (4), as

Faithfulness outranks Markedness completely, as in Tableaux **VI**.

VI	D ω [+atom] π [+auth][+part] ϕ		MAX(ω)	MAX(π)	ID(F)	*Indiv	*Index
$\boxed{\text{a}}$	D ω [+atom] π [+auth][+part] ϕ	(⇒1SG)	0	0	0	*	**
b	D ω [0atom] π [+auth][+part] ϕ	(⇒1PL)	0	0	*!	0	**
c	D ω [+atom] π [0auth][+part] ϕ	(⇒2SG)	0	0	*!	*	*
d	D ω [0atom] π [0auth][+part] ϕ	(⇒2PL)	0	0	**!	0	*
e	D ω [+atom] π [0auth][0part] ϕ	(⇒3SG)	0	0	**!	*	0
f	D ω [0atom] π [0auth][0part] ϕ	(⇒3PL)	0	0	***!	0	0

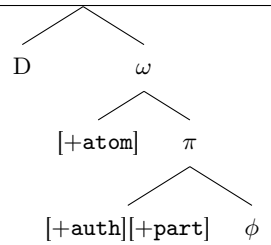
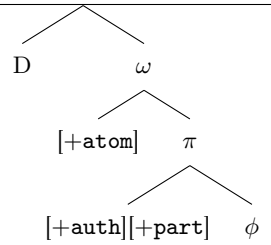
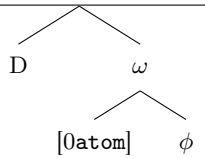
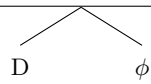
For candidate (e) to win, *Index must outrank Id(F), which in turn must outrank *Indiv, as in Tableaux **VII**.

VII 		Max(ω)	Max(π)	*Index	Id(F)	*Indiv
a 	(⇒1sg)	0	0	**!	0	*
b 	(⇒1pl)	0	0	**!	*	0
c 	(⇒2sg)	0	0	*!	*	*
d 	(⇒2pl)	0	0	*!	**	0
$\boxed{\text{e}}$ 	(⇒3sg)	0	0	0	**	*
f 	(⇒3pl)	0	0	0	***!	0

Hence, out of the non-harmonically-bound candidates (a, b, e, f), only (b, e, f) are possible proxies for singular polite self-reference. While (a) is certainly a possible outcome, it does not arise from a polite grammar.

What about the action of MAX(ω) and MAX(π)? So far, these constraints have been inactive because the candidates considered so far do not feature deletions of functional heads. However, inclusive and impersonal candidates do, represented in (b) and (c), respectively. Both (b) and (c) violate MAX(π) because the person head π has been deleted. (c) additionally violates MAX(ω) because the number head ω has been deleted as well. Note a difference in the violation profile of ID(F): while (b) violates ID(F) once because [+`atom`] in the input is changed to [0`atom`], (c) is impervious to ID(F) since deletion ensures that features are no longer in correspondence and does not violate ID(F).

For them to have a chance of winning, the markedness constraints *Index and *Indiv must be ranked above the faithfulness constraints MAX(ω) and MAX(π). If MAX(ω) additionally outranks ID(F), then inclusive (b) wins. This is a polite grammar: the language would rather delete the π projection completely than to have a positively specified π projection. This is illustrated in Tableaux **VIII**.

VIII 	*Index	*Indiv	MAX(ω)	ID(F)	MAX(π)
a  (⇒1SG)	**!	*!	0	0	0
$\boxed{\text{b}}$  (⇒INCL)	0	0	0	*	*
c  (⇒4)	0	0	*!	0	*

If ID(F) outranks MAX(ω) and MAX(π), then (c) wins. This is also a polite grammar: the language deletes the ω and π projections at all costs, since this results in a highly structurally impoverished structure. Recall that highly structurally impoverished structures are highly Taboo-compliant: as they lack functional projections, they preclude the possibility of being strengthened for the corresponding dimension of meaning. This is illustrated in Tableaux **IX**.

IX	D ω [+atom] π [+auth][+part] ϕ		*Index	*Indiv	ID(F)	MAX(ω) MAX(π)
a	D ω [+atom] π [+auth][+part] ϕ	(⇒1SG)	**!	*!	0	0 0
b	D ω [0atom] ϕ	(⇒INCL)	0	0	*!	0 *
$\boxed{\text{c}}$	D ϕ	(⇒4)	0	0	0	* *

5.1.1.1 Predictions *vs.* reality for singular polite self-reference

Summing up, then, almost the entire gamut of specificational combinations are possible outcomes for a 1SG input, with the exception of second persons, which are collectively harmonically bound. However, not all possible outcomes are *polite* outcomes, in the sense that they arise only from non-polite grammars. These predictions are summarized in (7), where the outcomes are described with their conventional names.

(7) (Im)possible and (non)polite outcomes for a 1SG input

Outcome	Possible?	Polite?	
1SG	Y	N	Faithness > Markedness
1PL	Y	Y	*Indiv > ID(F), *Index
2SG	N	—	collectively harmonically bound
2PL	N	—	collectively harmonically bound
INCL	Y	Y	MAX(ω) > ID(F), MAX(π)
3SG	Y	Y	*Index > ID(F) > *Indiv
3PL	Y	Y	*Index, *Indiv > ID(F)
4	Y	Y	ID(F) > MAX(ω), MAX(π)

Checking this set of outcomes against the empirical findings about this proxand from Table 2.8, repeated below, we correctly predict that 1PL, 1INCL, and impersonal are attested proxies of singular polite self-reference, and correctly predict that 2SG and 2PL are unattested. While 1SG is a possible outcome, it is not a polite one.

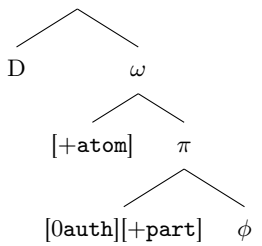
Recruitment asymmetries: Singular polite self-reference								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	6	0	0	4	0	0	1
Language	–	Persian	–	–	Guarijio	–	–	E. Kayah
		mā			*remé*			*ʔū*

We have also predicted that 3SG and 3PL are possible—and polite—outcomes. However, due to the already-small dataset ($n \approx 10$), we cannot validate this prediction with any certainty.

5.1.2 Deriving proxies of singular polite address

Singular polite address is the proxand that is most frequent, featuring the widest range of proxies. The input is that of 2SG, where π is less positively specified than it was for first person, featuring one (instead of two) positive specifications there, as [0`author`] and [+`participant`].

(8)



(⇒2SG)

In Tableaux **XI** below, (a) is the faithful candidate, featuring no changes of feature specifications. Candidate (b) is minimally different in terms of ω, involving a change from [+`atomic`] to [0`atomic`]. Candidates (c, d) are different in terms of [`author`], so that [0`author`] in the input is changed to [+`author`]. (d) involves an additional change from [+`atomic`] to [0`atomic`].

We first see two instances of harmonic bounding: (c) is harmonically bound by (a), and (d) is harmonically bound by (b). These are welcome predictions: since (c, d) involve a change from [0`author`] to [+`author`], they will necessarily contribute a stronger person presupposition. Indeed, proxies for singular polite address never take the shape of first person.

This leaves (a) and (b) in the competition. When *Indiv outranks ID(F), (b) wins over (a). We will see that this is the general constraint ranking that produces a plural proxy, being consistent with the ranking that produced a 1PL proxy from a 1SG input in the previous section.

			MAX(ω)	MAX(π)	*Indiv	ID(F)	*Index
XI	D ω [+atom] π [0auth][+part] ϕ						
a	D ω [+atom] π [0auth][+part] ϕ	(⇒2SG)	0	0	*!	0	*
$\boxed{\text{b}}$	D ω [0atom] π [0auth][+part] ϕ	(⇒2PL)	0	0	0	*	*
c	D ω [+atom] π [+auth][+part] ϕ	(⇒1SG)	0	0	*!	*	**
d	D ω [0atom] π [+auth][+part] ϕ	(⇒1PL)	0	0	0	**!	**

Of course, when all Faithfulness constraints outrank all Markedness constraints, the faithful candidate (a) wins. This is just to flag that this is a possible outcome; indeed; plenty of languages use 2SG for singular address without any reflexes of recruitment. Again, however, this is not a polite-making outcome.

Now let us additionally consider two groups of candidates in Tableaux **XII**. Candidates which change person feature specifications so that π is specified [0author][0part] are candidates (b) and (c) below, incurring one and two violations of ID(F), respectively. Candidates which delete π and/or ω functional projections are candidates (d) and (e), with both incurring a violation of MAX(π), and (e) incurring an additional violation of MAX(ω). We first observe that no candidate is harmonically bound, so that all candidates are able to surface with particular constraint rankings, which we will consider in turn now.

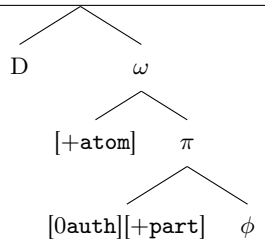
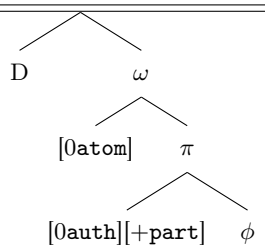
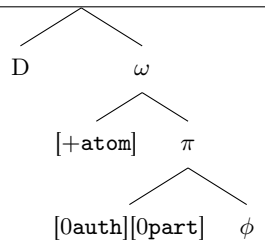
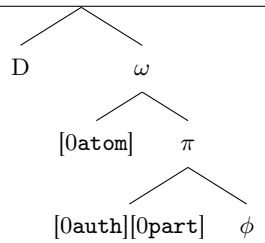
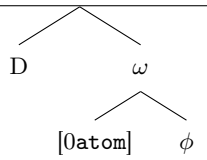
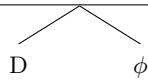
Under *Index > ID(F) > *Indiv, we see that (b) wins. This is the general constraint ranking that produces a third person singular proxy, being consistent with the ranking that produced the same candidate as a winner (although unattested) from a 1SG input above.

XII	D ω [+atom] π [0auth][+part] ϕ		MAX(ω)	MAX(π)	*Index	ID(F)	*Indiv
a	D ω [0atom] π [0auth][+part] ϕ	(⇒2PL)	0	0	*!	*	0
$\boxed{\text{b}}$	D ω [+atom] π [0auth][0part] ϕ	(⇒3SG)	0	0	0	*	*
c	D ω [0atom] π [0auth][0part] ϕ	(⇒3PL)	0	0	0	**!	0
d	D ω [0atom] ϕ	(⇒INCL)	0	*!	0	*	0
e	D ϕ	(⇒4)	*!	*	0	0	0

Under a minimally different ranking where *Index and *Indiv outrank ID(F), we see that (c) wins. This is displayed in Tableaux **XIII**. This is the general constraint ranking that produces a third person plural proxy, where the language would rather tolerate changes (two deletions, in this case) of feature specifications than use an indexical ([+atom]) or an individuator ([+part]).

XIII		MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
	D, ω; [+atom], π; [0auth][+part], ϕ					
a	D, ω; [0atom], π; [0auth][+part], ϕ (⇒2PL)	0	0	*!	0	*
b	D, ω; [+atom], π; [0auth][0part], ϕ (⇒3SG)	0	0	0	*!	*
$\boxed{\text{c}}$	D, ω; [0atom], π; [0auth][0part], ϕ (⇒3PL)	0	0	0	0	**
d	D, ω; [0atom], ϕ (⇒INCL)	0	*!	0	0	*
e	D, ϕ (⇒4)	*!	*	0	0	0

Now, let us consider outcomes when MAX(ω) and MAX(π) are not top-ranked so that the structurally impoverished candidates (d) and (e) have a fighting chance. This is displayed in Tableaux **XIV**. When MAX(ω) outranks ID(F), then (d) wins. This is a polite grammar: the language would rather delete the π projection completely than to have a positively specified π projection, and is consistent with a ranking that produces a inclusive winner, akin to what we have seen for a 1SG input above. Furthermore, ID(F) must outrank MAX(π) to ensure that (c) does not win over (d): this indicates that the language would rather change a feature specification than to make the more drastic move of deleting the π projection.

XIV 		*Index	*Indiv	MAX(ω)	ID(F)	MAX(π)
a 	(⇒2PL)	*!	0	0	*	0
b 	(⇒3SG)	0	*!	0	*	0
c 	(⇒3PL)	0	0	0	**!	0
$\boxed{\text{d}}$ 	(⇒INCL)	0	0	0	*	*
e 	(⇒4)	0	0	*!	0	*

When ID(F) outranks MAX(ω) and MAX(π), as in Tableaux **XV**, candidate (e) wins. This is also a polite grammar, and a very polite one: by deleting the ω and π projections at all costs, the resulting structure is highly structurally impoverished and it can never be strengthened for either number or person.

			*Index	*Indiv	ID(F)	MAX(ω)	MAX(π)
XV	D, ω; [+atom], π; [0auth][+part], ϕ		*Index	*Indiv	ID(F)	MAX(ω)	MAX(π)
a	D, ω; [0atom], π; [0auth][+part], ϕ	(⇒2PL)	*!	0	*	0	0
b	D, ω; [+atom], π; [0auth][0part], ϕ	(⇒3SG)	0	*!	*	0	0
c	D, ω; [0atom], π; [0auth][0part], ϕ	(⇒3PL)	0	0	**!	0	0
d	D, ω; [0atom], ϕ	(⇒INCL)	0	0	*!	0	*
$\boxed{\text{e}}$	D, ϕ	(⇒4)	0	0	0	*	*

5.1.2.1 Predictions *vs.* reality for singular polite address

Many feature-specificational combinations are possible outcomes for a 2SG input. Notable exceptions are [+auth][+part] forms, first person forms, which are harmonically bound. All other possible outcomes are also *polite* outcomes, with the exception of the faithful 2SG output. This is summarized in (9).

(9) (Im)possible and (non)polite outcomes for a 2SG input

Outcome	Possible?	Polite?	
1SG	N	—	harmonically bound
1PL	N	—	harmonically bound
2SG	Y	N	Faithfulness > Markedness
2PL	Y	Y	*Indiv > ID(F), *Index
INCL	Y	Y	MAX(ω) > ID(F) > MAX(π)
3SG	Y	Y	*Index > ID(F) > *Indiv
3PL	Y	Y	*Index, *Indiv > ID(F)
4	Y	Y	ID(F) > MAX(ω), MAX(π)

This aligns precisely with the empirical findings from Chapter 2, correctly predicting all and only the attested proxies of singular polite address, repeated from Table 2.6 below. We correctly predict that 2PL, INCL, 3SG, 3PL and impersonal are attested proxies of politeness, and correctly predict that 1SG and 1PL are unattested proxies. While 2SG is a possible outcome, it is not a polite one because it is maximally faithful.

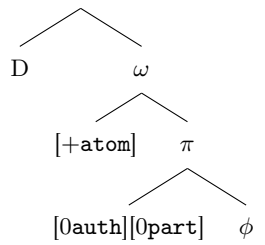
Recruitment asymmetries: Singular polite address								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	125	27	12	7	29
Language	–	–	–	French	Fehan Tetun	Italian	German	Ainu
				vous	*ita*	*Lei*	*Sie*	*an*

It is significant that the set of polite-making constraint rankings stays consistent across the two different inputs we have considered: certain constraint rankings yield certain proxies. A plural proxy arises from *Indiv > ID(F), while a third person proxy arises from *Index > ID(F). We will further illustrate this consistency for another proxand, that of singular polite reference.

5.1.3 Deriving proxies of singular polite reference

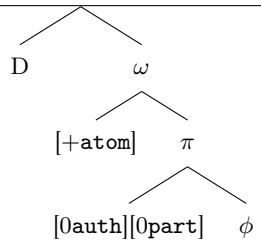
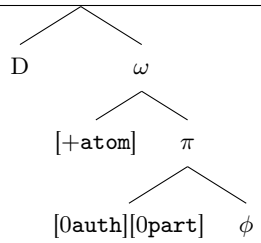
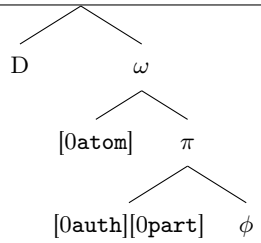
For singular polite reference, the input is that of 3SG, where π is the least positively specified: [0`author`] and [0`participant`].

(10)



(⇒3SG)

First, let us reassure ourselves that the ranking resulting in a plural proxy remains *Indiv > ID(F). 3SG (a) is the faithful candidate, and 3PL (b) involves minimally changing [+`atom`] to [0`atom`]. When *Indiv outranks ID(F), 3PL (b) wins. This is indeed the general constraint ranking that produces a plural proxy. (As always, the faithful candidate 3SG (a) wins when Faithfulness outranks Markedness.) This is displayed in Tableaux **XVI**.

XVI 	MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
a  (⇒3SG)	0	0	0	*!	0
$\boxed{\text{b}}$  (⇒3PL)	0	0	0	0	*

Now, let us consider candidates which are more positively-specified for person: these are maximally positively-specified [+`auth`][+`part`] candidates in (c) and (d), and [+`auth`][0`part`] candidates in (e) and (f). This is displayed in Tableaux **XVII**. Candidate elimination here proceeds rather rapidly, as each of (c, d, e, f) are harmonically bound: (c) by (e), (d) by (f), (e) by (a, b), and (f) by (e). Hence, (c, d, e, f) are predicted to never arise. Indeed, first and second person forms are unattested proxies for singular polite reference.

			MAX(ω)	MAX(π)	*Index	*Indiv	ID(F)
XVII	D, ω; [+atom], π; [0auth][0part], ϕ						
a	D, ω; [+atom], π; [0auth][0part], ϕ	(⇒3SG)	0	0	0	*!	0
$\boxed{\text{b}}$	D, ω; [0atom], π; [0auth][0part], ϕ	(⇒3PL)	0	0	0	0	*
c	D, ω; [+atom], π; [+auth][+part], ϕ	(⇒1SG)	0	0	**!	*	**
d	D, ω; [0atom], π; [+auth][+part], ϕ	(⇒1PL)	0	0	**!	0	***
e	D, ω; [+atom], π; [0auth][+part], ϕ	(⇒2SG)	0	0	*!	*	*
f	D, ω; [0atom], π; [0auth][+part], ϕ	(⇒2PL)	0	0	*!	0	**

Now, let us consider candidates which are structurally impoverished, as candidates (b) and (c) are in Tableaux **XVIII**. below. Immediately, we see that (b) is harmonically bound by the faithful candidate (a); this correctly predicts that the inclusive is never a proxy for singular polite reference. Furthermore, when ID(F) outranks MAX(ω) and MAX(π), candidate (c) wins. This is a polite grammar: the language would rather delete the π and ω projections at all costs, a pattern consistent with what we have seen for both 1SG and 2SG inputs from previous sections.

XVIII		*Index	*Indiv	ID(F)	MAX(ω)	MAX(π)
	D, ω; [+atom], π; [0auth][0part], ϕ					
a	D, ω; [0atom], π; [0auth][0part], ϕ (⇒3PL)	0	0	*!	0	0
b	D, ω; [0atom], ϕ (⇒1INCL)	0	0	*!	0	*
$\boxed{\text{c}}$	D, ϕ (⇒4)	0	0	0	*	*

5.1.3.1 Predictions *vs.* reality for singular polite reference

Fewer specificational combinations are possible polite outcomes for a 3SG input because it is already very presupposition-poor, owing to no positive specification at π. The OT analysis explains this, as many candidates are harmonically bound and can never arise under any constraint ranking. As for those candidates which are not harmonically bound, only 3PL and the impersonal arise from polite grammars. This is summarized in (11).

(11) (Im)possible and (non)polite outcomes for a 3SG input

Outcome	Possible?	Polite?	
1SG	N	—	harmonically bound
1PL	N	—	harmonically bound
2SG	N	—	harmonically bound
2PL	N	—	harmonically bound
INCL	N	—	harmonically bound
3SG	Y	N	Faithfulness > Markedness
3PL	Y	Y	*Indiv > ID(F)
4	Y	Y	ID(F) > MAX(ω), MAX(π)

This aligns precisely with the empirical findings from Chapter 2, correctly predicting all and only the attested proxies of singular polite reference, repeated from Table 2.7 below. We correctly

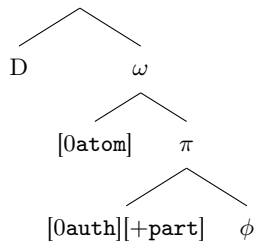
predict that 3PL and impersonal are attested, and correctly predict that 1SG, 1PL, 2SG, 2PL, and the inclusive are unattested. While 3SG is a possible outcome, it is not a polite one.

Recruitment asymmetries: Singular polite reference								
Recruit	1SG	1PL	2SG	2PL	1INCL	3SG	3PL	4
Frequency	0	0	0	0	0	0	28	8
Language	–	–	–	–	–	–	Wolaytta	W. Apache
							ʔetaa	*ko-*

5.1.4 Deriving proxies of plural polite address

For plural polite address, the input is a structure specified [0`author`], [+`participant`], and [0`atomic`].

(12)



(⇒2PL)

Consider alternatives which are also specified [0`author`][+`participant`] (candidates (a) and (b)), alternatives which are specified [+`author`][+`participant`] (candidates (c) and (d)), and structurally impoverished alternatives (candidates (e) and (f)). This is displayed in Tableaux **XIX**.

Candidate elimination proceeds rather quickly again with several instances of harmonic bounding. (b) is harmonically bound by (a), (c) is harmonically bound by (b), (d) is harmonically bound by (b), and (f) is harmonically bound by (e). This leaves (a) and (e) as the only possible outcomes. When we use the constraint ranking that has previously favored the inclusive—MAX(ω) > ID(F) and MAX(π)—as in the tableaux below, the inclusive (e) indeed wins.

XIX		*Indiv	*Index	Max(ω)	Id(F)	Max(π)
a	(D, ω([0atom], π([0auth][+part], ϕ))) (⇒2PL)	0	*!	0	0	0
b	(D, ω([+atom], π([0auth][+part], ϕ))) (⇒2SG)	*!	*	0	*	0
c	(D, ω([+atom], π([+auth][+part], ϕ))) (⇒1SG)	*!	**	0	**	0
d	(D, ω([0atom], π([+auth][+part], ϕ))) (⇒1PL)	0	**!	0	*	0
$\boxed{\text{e}}$	(D, ω([0atom], ϕ)) (⇒1INCL)	0	0	0	0	*
f	(D, ϕ) (⇒4)	0	0	*!	0	*

What about candidates specified [0`author`][0`participant`]? These are included as (b) and (c) in Tableaux **XX**. We see immediately that (b) is harmonically bound by (c). When *Index > Id(F), *Indiv; (c) wins, this is identical to the general constraint ranking favoring a proxy of higher person for previous inputs.

XX	D, ω; [0atom], π; [0auth][+part], ϕ		MAX(ω)	MAX(π)	*Index	ID(F)	*Indiv
a	D, ω; [0atom], π; [0auth][+part], ϕ	(⇒2PL)	0	0	*!	0	0
b	D, ω; [+atom], π; [0auth][0part], ϕ	(⇒3SG)	0	0	0	**	*!
$\boxed{\text{c}}$	D, ω; [0atom], π; [0auth][0part], ϕ	(⇒3PL)	0	0	0	*	0

5.1.4.1 Predictions *vs.* reality for plural polite address

Not many specificational combinations are polite outcomes for a 2PL input because it is relatively unrestricted already, possessing two out of three zero specifications. Furthermore, many candidates are impossible outcomes, being harmonically bound and never arising under any constraint ranking. As for those candidates not harmonically bound, only 3PL and the inclusive arise from polite grammars. This is summarized in (13).

(13) (Im)possible and (non)polite outcomes for a 2PL input

Outcome	Possible?	Polite?	
1SG	N	—	harmonically bound
1PL	N	—	harmonically bound
2SG	N	—	harmonically bound
2PL	Y	N	Faithness > Markedness
INCL	Y	Y	MAX(ω) > ID(F), MAX(π)
3SG	N	—	harmonically bound
3PL	Y	Y	*Index, *Indiv > ID(F)
4	N	—	harmonically bound

Empirical data concerning this proxand is extremely rare and further complicated with the presence of a morphological dual. However, §2.1.4.3 showed that a conjunction of higher number

and inclusivity is recruited for this proxand. In Limbu, for example, the 1INCL.DU is used for singular polite address, but the 1INCL.PL is used for two-or-more polite address, so that the inclusive plural is correctly predicted to be a polite outcome here, assuming that dualities are somewhat equivalent to pluralities[2]

As for the other possible polite outcome, 3PL, it is also correctly predicted to exist. In German, 3PL *Sie* is used for plural polite address (in addition to singular polite address).

5.1.5 Deriving politeness clines

The analytical advantage of OT lies in its ability to evaluate the suitability of several candidates relative to each other, providing a natural analysis for §2.2 which established the following clines for singular polite address. Recall that some recruits are more polite than others: the cline in (14) was informed by Indo-Aryan systems, while the cline in (15) was informed by languages which make clusivity distinctions. They establish two universal trends: first, third persons are recruited for higher tiers of politeness than the plural; second, the inclusive and the impersonal are recruited for the highest tiers of politeness.

(14) Singular polite address cline from Eastern Indo-Aryan

Recruit for...	2PL	◁	3SG	◁	3PL,	4
	2H pronoun, agreement		2MH pronoun, agreement		2HH agreement	2HH pronoun

(15) Singular polite address cline with inclusive

Recruit for...	2PL	◁	1INCL
	Marathi *tumhī*		Marathi *āpaṅ*
	Tukang Besi *ikomiu*		Tukang Besi *ikita*

These trends are explained by a *meta-ranking* of the various polite grammars proposed, as indicated below. The plural-favoring ranking (16a) is used for the lowest tier of politeness, the third person singular-favoring (16b) is used for an intermediate tier of politeness. For the highest tier of politeness, (16c) is used, which may produce third person plural, the inclusive, or the impersonal. While three different constraint rankings are contained in (16c), (16c) itself is one tier of politeness. The relevant meta-ranking is that *the lower Faithfulness is ranked or the*

[2]This is merely an assumption, since our privative feature system in §4.1 does not even touch on the issue of what kinds of LFs are responsible for deriving dual meanings. In fact, this entire dissertation pretends that duals, or finer-grained multiplicities, do not exist. This is because the semantic defaulthood of duals is not yet established, and it was decided at some point that it was not this dissertation's job to establish it. Interested readers are referred to Dvořák & Sauerland (2006) which discuss the semantic markedness of the Slovenian dual. They are also invited to recall Caddo, a language with the following approximation profile. The impersonal is recruited for singular polite reference, repeated as (ia), but the plural *-wa-* is recruited for dual polite reference in lieu of dual *-wiht-*, repeated from (41) as (ib).

(i) a. *dikat- **yi-** 'adih- hah? (=dikadiidihah)*
what- **IMPERS-** go- HAB
'Where is he/she (POLITE) going?'
Lit.: Where is <u>one</u> going?'

b. *kúyt- 'a= **wa**= dih- hah? (=kúyt'awadihah)*
where- AGENT= **PL**= go- HAB
'Where are them two (POLITE) going?'
Lit.: 'Where are <u>they (more than two)</u> going?' (Chafe 1990: 65-6, adapted)

The proxand for (ib) would be a LF that corresponds to 3DU in its conventional use. However, since this dissertation remains agnostic about the representation of duals, we do not discuss this issue further here.

higher Markedness is ranked, the more polite a grammar is. This is because when Markedness is ranked above Faithfulness, this is basically a state of affairs in which the Taboo has more s(w)ay over MP!.

(16) Meta-ranking of polite grammars

a. Proxy = plural when MAX(π), MAX(ω) > *Indiv > ID(F), *Index ***least polite***

b. Proxy = 3SG when MAX(π), MAX(ω) > *Index > ID(F), *Indiv ↑

c. Proxy = 3PL when MAX(π), MAX(ω) > *Indiv, *Index > ID(F)
Proxy = inclusive when *Indiv, *Index > MAX(ω) > ID(F), MAX(π)
Proxy = impersonal when *Indiv, *Index > ID(F) > MAX(ω), MAX(π) ↓
most polite

First, the escalation in politeness from (16a) to (16b) is due to the following reason. Third persons are recruited for higher tiers of politeness than the plural because violating *Index (to favor a less specified indexical; third person) is a more polite move than violating *Indiv (to favor a less specified non-indexical; plural). *Ceteris paribus*, a non-indexical category (number) has less referential power than an indexical category (person), as mentioned. Hence, (16b) is a more polite grammar than (16a) is.

Second, the escalation in politeness from (16b) to (16c) is straightforward. (16b) is a grammar in which the Faithfulness constraint ID(F) is equally low-ranked with a Markedness constraint *Indiv, while (16c) is a grammar in which the Faithfulness constraint ID(F) is the lowest-ranked single constraint. Hence, third person plurals (person- and number-neutral) are recruited for higher tiers of politeness than the third person singular (person-neutral without being number-neutral).

Third, the constraint rankings of (16c) which favor the structurally impoverished forms (inclusive, impersonal) are recruited for the highest tiers of politeness due to the following reasons. Violating MAX(π) and/or MAX(ω) is the most polite move because deleting π and/or ω projections completely precludes the possibilities of being later strengthened for person and/or number, making these alternatives the least informative in any context.

Note that 3PL and impersonal outcomes have no distinction between them in terms of their politeness relative to one another. Both are situated on the highest tier of politeness, as (14) indicates. This is rightly so: both these outcomes are person-neutral and number-neutral, albeit taking different routes to this very unrestrictive meaning. While the 3PL proxy arrives there by possessing only zero-specified features which are left unstrengthened, the impersonal arrives there by hosting neither π nor ω projection.

Lastly, we arrive at the question of switching between (non-)polite grammars. A single constraint ranking cannot be a language-wide setting: this would incorrectly predict the absence of tiered systems, that speakers have different registers at their disposal depending on their interlocutor and/or social setting. In principle, then, all polite grammars are available to all languages, but the choice of a polite grammar over another is subject to change even on an utterance-by-utterance basis. This is especially true when we consider mixed agreement, the phenomenon briefly discussed in the next section.

5.1.6 Deriving sentential politeness clines: a sketch for mixed agreement

The meta-ranking established in (16) above for pronominal politeness clines may be reflected in a separate politeness phenomenon, that of mixed agreement. *Mixed* agreement gets its name from featuring two types of agreement within the same sentence: *grammatical* agreement, which originates from the polite proxy of plural, and *notional* agreement, which reflects the cardinality

of the referent. Hence, in a language where the polite proxy of a language is set to plural, singular and plural agreement may be variably exponed throughout the sentence.

Consider French as an example. When the polite proxy, plural *vous*, is used politely for a singular, the verb obtains grammatical agreement (plural) throughout (17a-b). However, the adjective reflects notional agreement, reflecting the cardinality of the addressee. When *vous* is used politely towards one addressee, the adjective is singular (17a); used towards multiple addressees, the adjective is plural (17b).

(17) a. *Vous êtes loyal.*
2.HON COP.2PL loyal.**SG**
(politely towards a single addressee): 'You are loyal.'
b. *Vous êtes loyaux.*
2.HON COP.2PL loyal.**PL**
(politely or not towards multiple addressees): 'You are loyal.'

There is a fair amount of variation for where the split between grammatical and notional agreement occurs within the sentence. In Czech, for example, finite verbs show plural grammatical agreement, while the participle and predicative adjective show singular notional agreement (18). This phenomenon is fairly general across European languages and has been extensively discussed by Comrie (1975); Corbett (1983); Wechsler (2011); Wechsler & Hahm (2011); Arsenijević (2014); Despić (2017), and Puškar-Gallien (2019), among others. (18) displays variation in European languages (Comrie 1975: 410); an asterisk indicates dialectal variation.

(18) *Vy jste byl-a dobr-á.*
2**PL** AUX.2**PL** been.FEM.**SG** good-FEM.**SG**
(politely towards a single female addressee): 'You were good.' (Comrie 1975: 408)

(19)

	Verb	Participle	Adjective	Noun
Polish	PL	SG/PL*	SG/PL*	SG
Serbo-Croatian	PL	PL	PL	SG
Russian	PL	PL	PL	SG
Czech	PL	SG	SG	SG
Modern Greek	PL	SG	SG	SG
Romanian	PL	SG/PL	SG	SG
French/Italian	PL	SG	SG	SG

However, here we focus on a subtype of mixed agreement which has been understudied in the relevant literature. I will term it *socially mixed agreement*, where a "mixed" construction is one of several available agreement options which carries an additional layer of social meaning. In what follows, I will illustrate instances of socially mixed agreement, distinguishing it from the aforementioned subtype which I will term conventionally mixed agreement.

Modern Slovenian is a language with socially mixed agreement: three constructions are possible for singular address. (20a), with uniform plural agreement, is the default for polite singular address. (20b), with uniform singular agreement, is the default for informal singular address. Crucially, (20c), features mixed agreement: grammatical plural agreement on the auxiliary, but notional singular agreement on the verb. (20c) is interpreted as a "half-polite" singular address, used in social contexts which have elements of both distance and intimacy; for instance, in a TV interview where interviewer and interviewee have met for the first time, but the interviewer does not want their conversation to be stilted and overly formal.

(20) a. *Vi s- te kavo naroč- ili.*
2**PL** AUX.2 **PL** coffee.F.ACC order- 2**PL**
b. *Ti s- i kavo naroč- u.*
2**SG** AUX.2 **SG** coffee.F.ACC order- 2**SG**
c. *Vi s- te kavo naroč- u.*
2**PL** AUX.2 **PL** coffee.F.ACC order- 2**SG**
(varying degrees of politeness towards a singular addressee): 'You ordered coffee.' (Mitrović 2025)

Another instance of socially mixed agreement can be found in Persian[3]. The proxy of politeness language-wide is plural, so that 2PL *shoma* is used for singular polite address in lieu of 2SG *to*. As in Slovenian, using plural *shoma* with uniform plural verbal agreement signals the most social distance, while using singular *to* with uniform singular verbal agreement signals the least social distance. It is the mixed construction which signals an intermediary amount of social distance as in (21a-b): the pronoun is plural, but the verb carries notional singular agreement.

(21) a. *mærjæm mi-goft* ***shoma*** *mæɣale-je biozenetik-ro dar-**i**.*
Maryam DUR-say.PST **2PL** article-EZ biogenetic-OM.Q have-**2SG**
'Maryam told me that you have this biogenetic article'.
b. *mi-g-æm* ***shoma*** *mæɣale-je zæbanshenasi-o dar-**i**...*
DUR-say.Q.1SG **2PL** article.EZ linguistics-OM.Q have-**2SG**...
'I say, do you have the linguistics article...?' (adapted from Saeli & Miller 2018: 173, 178)

For such mixed constructions, Saeli & Miller (2018: 174) note that "[w]hereas the use of the singular verb signals higher sociocultural comfort, the incorporation of the plural pronoun still functions as a creator of social distance". This accords with earlier sociolinguistic work by Baumgartner (1982) on Persian, which highlights the systematic use of mixed agreement constructions. For instance, a speaker would use the plural pronoun towards elder addressees: uncles, aunts, and in-laws on one hand, and parents on another hand. However, the speaker would use plural agreement towards the former, and mixed agreement towards the latter. In this sense, the mixed construction was used to signal both intimacy (from addressing close kin) and formality (from addressing elders). Baumgartner's results also show spontaneous switches between constructions occurring during conversations that corresponded to the level of intimacy between interlocutors. This may be taken as evidence for our pragmatic proposal of politeness as a *synchronic*, on-line process of evaluating grammatical alternatives.

Similarly, Jensen (1981: 59) reports 14 ways of saying the same utterance in Portuguese, where the pattern of agreement (whether it is uniform plural agreement, uniform singular agreement, or mixed agreement) carries different socio-pragmatic connotations.

Hence, we observe that in languages with socially mixed agreement, more grammatical agreement indicates more politeness (and equivalently, more notional agreement indicates less politeness). In this way, languages like Persian are akin to the multiple-tier politeness systems we saw for pronouns in §2.2, except that here, the tiers of politeness are distinguished by verbal agreement, and not by the recruitment patterns in pronouns.

Analytically, there are three possibilities for handling mixed agreement (social and non-social): syntactic, semantic, and a hybrid between the two.

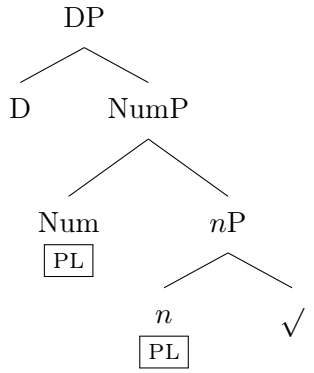
The hybrid approach is the most established in the literature. It assumes a distinction between *index v. concord* ϕ-features (Wechsler & Zlatic 2003; Wechsler & Hahm 2011; Wechsler

[3]Thank you to Amir Anvari for bringing this to my attention.

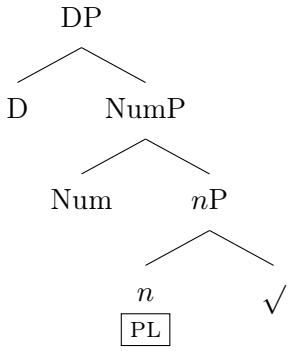
2011; Ackema 2014). Concord features encode syntactic properties, while index features encode semantic properties and are tied to referential indices. Pronouns contain both concord and index features, which variably interact with different agreement targets in the sentence. In a mixed construction, the grammatical plural agreement would be the result of agreement with a concord feature, while notional singular agreement the result of agreement with an index feature. However, this approach is untenable in the current account, which, fortunately or unfortunately, only posits one (semantic) type of ϕ-feature.

Sinha (2025) uses a purely syntactic approach, which posits number as a uniform feature situated on two distinct heads, Num^0 and n^0. The number system is privative, where singular is modeled as the absence of PL (as per Nevins 2011). Regular plurals have a PL feature sitting on both Num^0 and n^0 (22), while plural proxies of politeness only have a PL feature sitting on n^0 and not Num^0 (23). This difference is used to derive the possibility of singular agreement with the latter but not the former. Probes vary in their ability to access the PL feature on n^0. Plural agreement results when it is accessed, while singular agreement results when this access is blocked. The variable access is made possible by the Interaction-Satisfaction model of agreement (Deal 2015).

(22)



(23)



Adapted for the current account, there might be a Faithfulness constraint mandating against a change of number specification on Num^0 from singular to plural, *i.e.* whether the plural proxy of politeness in (23) turns into a regular plural as in (22).

However, this account relies on plural being the marked number value, an assumption fundamentally incompatible with what we have established in Chapter 3: that plural is the semantic default for number, not singular. As a remedy, one could posit that plurals are morphosyntactically marked but semantically unmarked; this would avoid a conflict in underlying assumptions. However, such a remedy seems inadequate. While plural seems to be morphosyntactically marked in nouns, the same does not seem to pertain for pronouns (e.g. Greenberg's Universal 45: "If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also.")

This leaves us with a purely semantico-pragmatic approach, the analytical option that aligns best with the current account. Let us understand "notional" agreement as the result of a constraint ranking with higher-ranked Faithfulness constraints (which reflect aspects of the context), while "grammatical" agreement as the result of a constraint ranking with higher-ranked Markedness constraints (which are only operative in polite contexts anyway). Then, the proposed meta-ranking straightforwardly explains the patterns of socially mixed agreement: *the lower Faithfulness is ranked or the higher Markedness is ranked, the more polite a grammar is.* I do not offer any insight about how the number feature is transmitted to adjectives and participles. What is crucial here is only whether concord on these predicates is exponed, and not whether concord pertains in the first place (it does).

Of all three approaches to mixed agreement, the last approach may have the advantage of being able to explain *why* the phenomenon pertains in the first place. Mixed agreement may be a way of addressing the inherent tension produced by recruitment: the tension between grammatical reference (dictated by the negative politeness/Markedness) and notional reference (dictated by aspects of the context/Faithfulness). As we have repeatedly seen, recruitment creates ambiguity about whether the grammatical or social meaning of *vous* is intended. Of course, such ambiguity is easily resolved by being in the conversational context and ascertaining whether it is a polite context or not. However, if aspects of the context are not readily available, (24a) in a non-polite context is a violation of Quality ("Be truthful, and do not give information that is false or that is not supported by evidence").

(24) *Avez vous le livre?*
have.PRES.2**PL** 2**PL** the book
a. **grammatical**: 'Do <u>you (plural)</u> have the book?'
b. **social**: 'Do <u>you (singular + polite)</u> have the book?'

Mixed agreement would be a way for the grammar to resolve such ambiguities when aspects of the context are not readily available. Certain corners of the grammar are offered up to reflect aspects of the context, such as the adjective in (25).

(25) *Vous êtes loyal.*
2.HON COP.2PL loyal.**SG**
(politely towards a single addressee): 'You are loyal.'

Other than mixed agreement, languages may address the same tension via *repluralization*. This is a process which occurs in the wake of recruitment: when the plural is recruited for polite singular, this creates a plural/polite syncretism but levels the plural/singular distinction. Hence, languages may innovate a new plural form to restore the plural/singular distinction. Such a language is Basque (26): after plural *zu* was recruited for politeness, the nominal plural morpheme *-ek* was used for repluralization, forming the current plural *zu-ek*. (For more discussion of this, see Wang 2024).

(26) Basque repluralization (second persons)

	SG	PL
plain	hi	zu
polite	zu	zu

→

	SG	PL
plain	hi	zu-ek
polite	zu	zu-ek

Mixed agreement and repluralization, then, are both processes that preserve plural meaning after the plural pronoun has been co-opted for singular honorification. It is interesting that none of the languages exhibiting mixed agreement also display repluralization in the pronouns. Basque, for instance, has a very rich paradigm of verbal agreement, but does not display mixed agreement. This suggests that these two processes are merely different ways of obeying the same Faithfulness pressure.

However, this approach requires much, much more elaboration. The adapted OT mechanism would be evaluating inputs and outputs of a fundamentally different kind: not DPs, but rather (components of) entire sentences. How such pragmatic optimization is implemented at the sentence level remains to be seen. Additional assumptions must also be made in order to correctly derive variation in terms of which sub-sentential components can be targeted by Markedness.

Having flagged these options, I leave the specific implementation of mixed agreement for future work.

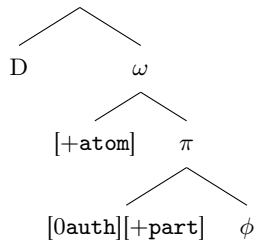
5.1.7 Summary

The OT analysis delivers on several desiderata. For polite address and reference, it delivers all and only their attested proxies. For polite self-reference, it correctly predicts the shapes of the attested proxies, but predicts some other shapes to be attested. OT enables us to generalize different constraint rankings as distinct polite grammars. Regardless of the input, specific constraint rankings will favor plural, third person, inclusive, or impersonal proxies. This also enables us to refer to languages which *generally* indicate politeness with plural, such as Tamil, which only recruits the plural for politeness across all three proxands.

To summarize, what follows is a schematization of the possible proxies for polite address, comparing the relevant structures in polite v. conventional usages. I choose polite address to summarize because it has the widest range of attested proxies.

For an input like 2SG as in (27), its positive specifications [+`participant`] and [+`atomic`] are flagged by OT constraints *Index and *Individuate, respectively (correlating to the sub-Taboos on Presumptuousness).

(27)



(⇒2SG)

This starts the redressive process, which consists of selecting candidates that variably violate ID(F), MAX(ω), or MAX(π) to satisfy *Index and/or *Individuate, as in (28). Each redressive candidate may emerge via different constraint rankings that favors it as the winner, as the summary in (28) (adapted from (7)) shows.

(28) Redressive Candidates and Requisite Rankings

ⓐ number-neutral second person	*Indiv > ID(F), *Index
ⓑ person-neutral singular	*Index > ID(F) > *Indiv
ⓒ person- + number-neutral	*Index, *Indiv > ID(F)
ⓓ person- + number-neutral	MAX(ω) > ID(F) > MAX(π)
ⓔ person- + number-neutral	ID(F) > MAX(ω), MAX(π)

Within this pool of attested proxies, we are also able to explain the shape of politeness clines formed between multiple proxies. ⓐ (⇒2PL) is recruited for a lower tier of politeness than ⓑ (⇒3SG) is. This is because redressive candidate ⓑ avoids indexical reference (being [0`part`]) while redressive candidate ⓑ merely avoids individual reference (being [0`atomic`]). In addition, ⓑ (⇒3SG) is recruited for a lower tier of politeness than ⓒ, ⓓ, and ⓔ (⇒3PL, INCL, 4) are. This is because while redressive candidates ⓒ, ⓓ, and ⓔ are person- and number-neutral, redressive candidate ⓑ is only person-neutral.

Lastly, since ⓒ, ⓓ, and ⓔ (⇒3PL, INCL, 4) are all person- and number-neutral, they are never recruited for distinct tiers of politeness. If a language recruits either as part of a multiple-tiered system, these redressive candidates are correctly predicted to only appear in the highest tier of politeness, with no distinction in politeness between them.

In addition to the shapes of attested proxies, we also derived the shapes of unattested proxies. Those ϕ-featural shapes are either harmonically bound (never able to arise under any

constraint ranking) or arise only under non-polite grammars.

One thing we have not accounted for is whether languages decide to allow approximation for address, reference, self-reference, or a combination of these. Approximation is very rare for self-reference, but highly frequent for address. This variation in frequency may be functional: if politeness is a social game, it is best played with an immediately present other (second person), rather than with a non-immediately present other (third person) or with oneself (first person). This account does not offer any insight otherwise.

5.2 Other ϕ-featural categories: Gaps?

The analysis sketched is very general: it does not depend on the vagaries of being a number/person category itself, but on asymmetries in presuppositional strength. Other ϕ-featural categories, such as gender and spatial deixis, also exhibit asymmetries in presuppositional strength, and are thus predicted to be categories which participate in approximating politeness. However, we will see this is unattested for gender, and only marginally attested for spatial deixis.

5.2.1 Gender

[4]In languages with (at least) a feminine-masculine opposition, previous work has shown that gender does correlate with markedness oppositions (Bobaljik & Zocca 2011, Sauerland 2008b). In several European languages, it is the masculine that is less marked. This can be shown with the behavior of noun pairs like *actor/actress* under ellipsis. The masculine noun licenses an elided feminine noun, but not vice versa. This explains why (29a) is deemed more acceptable than (29b) for Brazilian Portuguese *ator/atriz*. (Elided constituents are bracketed in gray.)

(29) a. ?*O Paulo é ator e a Fernanda também é atr-iz.*
the Paulo is actor and the Fernanda also is actr-ess
'Paulo is an actor and Fernanda is too.'

b. #*A Fernanda é atr-iz e o Paulo também é ator.*
the Fernanda is actr-ess and the Paulo also is actor
Intended: 'Fernanda is an actress and Paulo is too.' (Bobaljik & Zocca 2011: 144-5)

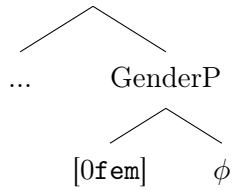
A similar asymmetry holds with English *actor/actress*:

(30) a. Froggy is a waiter, and Anne is a waiter too.
b. #Anne is a waitress, and Froggy is a waitress too.

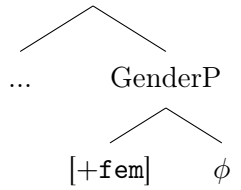
This suggests that gender might benefit from a privative analysis like we have sketched above, where the unmarked pair member (masculine) features zero specification (31a) and the marked pair member (feminine) positive specification (31b). (31a) is born gender-neutral, and only obtains its conventional meaning of masculine via competition with the feminine (31b).

[4]Much of this section is adapted from Wang 2023: §6.1.

(31) a.



b.



Yet, masculine-marked pronouns are never found to be recruited for female-directed politeness. For that matter, gender has absolutely no approximative reflex at all. Even though gender systems are set up variably across languages, adhering to distinctions other than masculine *v*. feminine, there is no evidence that any gender category is recruited. Either a privative analysis is on the wrong track, or approximation is not as general a phenomenon. Here I offer a few speculations, but leave this gap for future work.

To start, "gender" is far from a uniform category across languages. In Romance, many nouns may show a discrepancy between grammatical and biological values (e.g. Italian *il soprano* 'the.MASC soprano.MASC' denoting normally-female sopranos, Spanish *la persona* 'the.FEM person.FEM' denoting men, etc.). This shows that gender may be a purely formal reflex of noun categorization. Languages may also choose to make different types of gender distinctions, varying from a feminine-masculine opposition, to human-nonhuman, to animate-inanimate, and so on. In comparison, person and number are much more uniform categories, with languages partitioning the relevant ontological spaces relatively homogeneously (Harbour 2014, 2016).

Furthermore, gender is often parasitic on other phi-features: for example, gender is marked only in third person, preferably third person singular, in many languages. This observation is captured by Greenberg's (1963) Universal 36: "If a language has the category of gender, it always has the category of number." Another explanation may be the low frequency of gender marking in pronouns: across the combinations of number and person, gender marking is most frequent in 3SG pronouns. This is indicated by Greenberg's (1963: 96) Universals 44-45 (32a-b).

(32) a. If a language has gender distinctions in the first person, it always has gender distinctions in the second or third person or in both.
b. If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also.

Thus, we only expect to see gendered approximation for languages which allow approximation for singular polite reference, a proxand which is already relatively infrequent.

Lastly, I offer a further difference between gender and number/person, which forms part of a speculative explanation as to why gender never has honorific effects. The Taboo on Presumptuousness that has been proposed is the grammatical reflex of a negative politeness strategy. However, readers might wonder if there exists a grammatical reflex corresponding to *positive* politeness strategies: those which acknowledge that an interlocutor's wants and needs are respected and "desirable to at least some others".

I stipulate that the lack of honorific gender might be a relic of the pressure of addressing *positive* face-wants for the following reason. While one's cardinality is not generally considered an important part of one's social image, or face, one's gender generally is (at least in indigenous

communities where many of the languages considered in the current sample are spoken). Unlike number/person, gender is grounded in one's identity and one's presentation of this social identity. The use of honorific gender might be very threatening to positive face, since a gender mismatch ostensibly overlooks the interlocutor's own desires for how this important part of their identity is perceived. Indeed, in Khwe (Southwest Africa, Khoisan), addressing another using the reverse gender has pejorative effects. In (33a), a male addressee is referred to with the 2SG.MASC pronoun *tcá*, and later addressed pejoratively as a woman via the feminine vocative *-cè* in (33b).

(33) a. ***Tcá*** *tàn* *tiú* *ti* *múù è!*
2SG.**MASC** stand.up then 1SG see IMP
'Stand up, let me see you!'
b. *Ah, tàn-**cé!***
ah stand.up-2SG.**FEM**.VOC
'Ah, stand up, you "woman"!' (Kilian-Hatz 2008: 175-6)

5.2.2 Spatial deixis

Pre-theoretically, the most obvious reflex of marking "social distance" would be spatial distance, making it an obvious, if not highly frequent, proxy category. However, while "polite deixis" is attested, it is not very frequent; I list the attested instances below.

The most robust datapoint comes from Basque (isolate, Ignacio Hualde 2003: 151), which makes a threeway distinction of proximal *v.* mesial *v.* distal in its demonstratives. In some Eastern varieties, the formal second-person singular pronoun *ori* is derived from the mesial demonstrative *hori*. In Southern varieties, the pronoun used for the highest tier of polite address is *berori*, which involves the mesial demonstrative stem *-ori* (*ber-* is used to express reflexive "same" or intensive "again"). Lower tiers of polite address are obtained by recruiting the plural.

K'iche' (Mayan, Guatemala) boasts four distinctions in its demonstratives: proximal, medial, distal, and far distal (which is increasingly in disuse). Out of these, the distal demonstrative *lá* is the one used for honorification (Omer Preminger, p.c.).

Two further examples come from larger pronoun inventories, which this dissertation only briefly touches upon. In Malayalam (Dravidian, Asher & Kumari 1997: 259), the polite forms of address *aviʈunnə* and *aŋŋunnə* are derived from the distal demonstrative *aŋŋə* 'there'. These forms, literally 'from there', are becoming obsolete.

Japanese *anata*, one of polite address forms, presents a particularly interesting case because there is also evidence for the distal demonstrative being semantically unmarked. Whole books have been written about *anata* alone and its use conditions (see Yonezawa 2021), but I will focus only on its diachrony here. *anata* was derived from distal *a-*, the former linker *na-* (now *no-*), and *ta*, the former term for 'direction, side, place'. Hence, this form was literally a deictic marker "that direction, that side (far away)" that was recruited for polite address. *a-* is very productive in the current language, being also present in *a-itsu* 'that guy (derogative)' and *a-no* 'that (demonstrative)'. Its medial counterpart is *ko-* while its proximal counterpart is *so-*; these deictic categories do not seem to be involved in forming polite address terms. Interestingly, distal *a-* seems to be the semantically unmarked deictic value[5]: *ano* 'that', not *kono* or *sono*, is the demonstrative used to refer to something in previous discourse, even though there is no sense in which discourse objects are always "distal"[6].

As with gender, an explanation for the marginality of deictic approximation may be the low

[5]Thanks to Shigeru Miyagawa for pointing this out to me.

[6]A similar fact pertains in English: the distal demonstrative has a wider range of uses than the proximal one, so that it is the distal *that* that is reused for heading relative and complement clauses.

frequency of deictic marking in pronouns, which, as far as I know, is only made in the third person. Thus, we would only expect to see deictic approximation for languages which allow approximation for singular polite reference in the first place, explaining their rarity.

5.3 Comparison to other analyses

[7]Previous analyses of honorific pronouns typically assume a feature specialized for honorification: for example, Simon (2003) and Ackema & Neeleman (2018) assume [HON(ORIFIC)] in their representations of honorific pronouns. Macaulay (2015a) takes a similar stance, assuming the feature [STATUS]; with Portner, Pak & Zanuttini (2019) proposing [`formal`]. Corbett (2012) takes a more cautious approach, considering a [HON] feature only for a handful of languages. Others do not situate their honorific feature in the morphosyntax, but in the expressive dimension within a multidimensional semantics (e.g. Potts 2005, McCready 2019). Non-generativist perspectives (e.g. Listen 1999: 44) use conceptual metaphors such as POWERFUL IS PLURAL to capture honorific uses of plural.

Here, I review two recent proposals in detail: the impoverishment-based proposal of Ackema & Neeleman (2018), and the agreement-based proposal of Portner, Pak & Zanuttini (2019). Both assume a grammatical feature dedicated to honorification, and are summarized in §5.3.1. They are critiqued in light of the typological restrictions on recruitment established in Chapter 2 (§5.3.2), and the impoverishment mechanism in particular is shown to overgenerate, wrongly predicting unattested patterns (§5.3.3).

5.3.1 A dedicated grammatical feature for honorification

Here, we summarize the impoverishment-based proposal of Ackema & Neeleman (2018), and the agreement-based one of Portner, Pak & Zanuttini (2019) in turn.

For Ackema & Neeleman (2018) (henceforth A&N), the grammatical feature [HON] is included in the representation of all honorific pronouns, formally distinguishing plain from honorific pronouns. [HON] is on par with other pronominal features, such as those of number and person. However, [HON] does not affect reference, merely indicating that the members in its set are honorable. (A&N develop their own features for person and number, but their analysis of honorifics is fairly theory-neutral.)

As we have extensively seen, a hallmark of an honorific pronoun is the mismatch it creates. A&N explain such mismatches with *impoverishment*, where [HON] conditions deletion of the offending feature, either at LF or PF. Consider how this works for, say, French honorific *vous*, which exhibits a number mismatch. A&N assume that, whereas plain *vous* contains features for a second person plural denotation, honorific *vous* additionally contains a [HON] feature. In honorific contexts, [HON] triggers deletion of the plural feature. This is termed *LF impoverishment*, which conditions deletion of features after syntax but before interpretation, as in (34).

(34) PL → Ø / [_ HON] (at LF)

The resulting pronoun is number-neutral, concordant with the interpretation of honorific *vous*: it can be used respectfully towards one or more addressees. This can be seen from variable number agreement on the adjective. If *vous* is respectfully directed towards one addressee, adjectival agreement is singular (35a). Towards multiple addressees, adjectival agreement is plural (35b).

[7]Large portions of this section are taken from Wang (2023: §3).

(35) a. *Vous êtes* ***loyal.***
2.HON COP.2PL loyal.**SG**
'You (POLITE) are loyal.' (honorific singular address)
b. *Vous êtes* ***loyaux.***
2.HON COP.2PL loyal.**PL**
'You all (POLITE) are loyal.' (honorific plural address)

This also explains why honorific *vous* obligatorily triggers plural verbal agreement, as this impoverishment happens at LF, after syntactic agreement has taken place. The variable adjectival agreement is taken to be semantic agreement (see Ackema 2014, also Wechsler (2011) on notional agreement).

Consider further how this would work for German honorific *Sie*, which creates both number and person mismatches when deployed for honorific address. Whereas plain *sie* contains features for a third person plural denotation, honorific *Sie* additionally contains a [HON] feature. In this case, impoverishment takes place at PF. [HON] triggers impoverishment of the feature responsible for a second person interpretation (in their framework, the [PROX(IMAL)] feature) at PF (36a), so the resulting pronoun is phonologically identical to the third person plural pronoun *sie*. This eliminates the person mismatch. Another dose of impoverishment, this time at LF (36b), derives the desired number-neutrality, eliminating the number mismatch.

(36) a. PROX → Ø / [_ DIST HON PL] (at PF)
b. PL → Ø / [_ HON] (at LF)

PF impoverishment allows honorific *Sie* to lead a double life: it is phonologically identical to third person plural, but syntactically identical to a second person plural. PF impoverishment only establishes a surface similarity to 3PL *sie*: the authors cite Simon (2003) in observing that honorific *Sie* syntactically behaves like the second person plural pronoun, *ihr*, rather than the third person plural pronoun, *sie*. Evidence is given from close appositions (37a-b) and relative clauses (38). In both constructions, the distribution of honorific *Sie* patterns with that of 2PL *ihr*, away from that of 3PL *sie*[8].

(37) a. ***ihr/Sie*** *Finnen*
2PL/HON Finns
'You (plural)/you (POLITE) Finns'
b. ****sie*** *Finnen*
3PL Finns
Intended: 'them Finns'

(38) a. ***ihr/Sie,*** *die* ****(ihr/Sie)*** *Bücher sammeln, besitzt fünf Ausgaben der Bibel.*
2PL/HON, who 2PL/HON books collect own five editions the.GEN Bible
'You (plural)/you (POLITE), who collect books, own five editions of the bible.'
b. *Sie, die (****sie****) Bücher sammeln, besitzt fünf Ausgaben der Bibel.*
3PL, who 3PL books collect own five editions the.GEN Bible

[8]Although, it remains a mystery under A&N's account as to why honorific *Sie* is always incompatible with second person plural verbal agreement (i).

(i) a. ****Seid*** ***Sie*** *müde?*
COP.**2PL HON** tired
Intended: 'Are you (POLITE) tired?'
b. ****Sie*** ***geht*** *zum Platz.*
HON go.**2PL** to.the square
Intended: 'You (POLITE) are going to the square.'

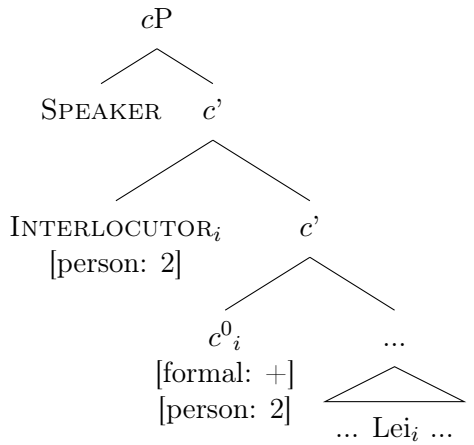
Intended: 'They, who collect books, own five editions of the bible.'

In sum, A&N's proposal of recruited honorifics rests on the assumption that [HON]-conditioned impoverishment (at LF or PF) resolves the mismatches that recruited honorifics create.

Similar in spirit to A&N's impoverishment analysis is Portner, Pak & Zanuttini (2019) (henceforth PPZ), who also propose a dedicated feature for honorification, [`formal`], sensitive to whether the interaction calls for formality or not. Like much previous work on speaker-addressee relations, components of the speech act are assumed to be syntactically encoded (e.g. Speas & Tenny 2003, Haegeman & Hill 2013) so that (null) abstract representations of SPEAKER and INTERLOCUTOR are made available by the syntax. [`formal`] originates on c^0, a high functional head above CP, so that c^0P is the syntactic locus for utterance-level meanings. [`formal`] can be copied onto bound pronouns in the domain of c^0 via an operation similar to Kratzer's (2009) Feature Transmission.

In this analysis, all second person pronouns are bound by INTERLOCUTOR via operator-variable agreement (the operator being INTERLOCUTOR, and the variable being the pronoun). c^0 is the functional head mediating this agreement. Having this agreement relationship with INTERLOCUTOR is what makes a pronoun second person, as PPZ assume that pronouns are minimally specified when they enter the derivation (*à la* Kratzer 2009). This configuration is illustrated in (39) with the Italian honorific pronoun *Lei*. The honorific, the INTERLOCUTOR head, and the c^0 head are all in an agreement relation involving a [`formal`] feature.

(39)



What determines the form of the honorific pronoun is a set of spell-out rules sensitive to the specification of the binary feature [`formal`]. When a pronoun is bound by INTERLOCUTOR, and c is valued [-`formal`], it is spelled out as *tu*. When a pronoun is bound by INTERLOCUTOR, and c is valued [+`formal`], it is spelled out as *Lei*. This is schematized in (40).

(40) a. (pronoun) → *tu* / INTERLOCUTOR$_i$ ... $c^0{}_{[\texttt{-formal}]}$ ··· _
b. (pronoun) → *Lei* / INTERLOCUTOR$_i$ ... $c^0{}_{[\texttt{+formal}]}$ ··· _

5.3.2 Typological inadequacy

How do these non-recruitment-based proposals account for the recruitment asymmetries established in Chapter 2? At first glance, it seems that there are no restrictions placed on whether a ϕ-featural value can carry [HON] or not, or when c^0 is valued [+`formal`] or not. While A&N posit some restrictions on which ϕ-featural values may be recruited for politeness, PPZ do not. Furthermore, neither account comprehensively derives the full gamut of typological patterns,

leaving unaccounted the recruitment asymmetries in single-tier politeness systems (§2.1.5) and in multiple-tier politeness systems (§2.2).

The typological restrictions that A&N derive stem from the assumption that certain ϕ-featural values can never be a target of impoverishment. For number features, the restriction on recruiting singular stems from a privative view of number, similar to the view adopted by this account. Plural forms are specified via a PL feature. The absence of a feature has interpretative effects, so that the absence of plural implies singular. Since singular is never specified in morphosyntax, it can never be the target of impoverishment. This explains why singular pronouns are never recruited for honorifics.

As to why first and second persons are never recruited for honorification, A&N give functional explanations. Second person singular pronouns are never recruited for honorific singular address, as the intended respectful effect would not be detectable via a mismatch. Neither are first person singular pronouns, due to learnability considerations: a mismatch is necessary for a pronoun to be interpreted as honorific. But a mismatch is never possible for the first singular, as it is always interpreted as the speaker, making its acquisition as a honorific "difficult if not impossible" (p. 46). Since honorification is definitionally equivalent to grammaticalized social meaning, this social meaning would be difficult to detect if there were no grammatical indications in the form of a mismatch. However, a functional explanation is not fully satisfactory because it undergenerates. A mismatch would be detectable if second person pronouns were recruited for honorific reference or honorific self-reference, but second person proxies for these rarer proxands are also unattested, as was seen in §2.1.4.1 and §2.1.4.2, respectively.

As for PPZ, the authors themselves state that their analysis "can't provide a detailed discussion of the mapping between the abstract features and the features on the [polite] pronouns" (p. 31). Thus, any relationship between proxy and proxand is denied. While this was never an analytical goal of their [`formal`]-based theory, leaving the typological trend unacknowledged is an unparsimonious move. As the analysis relies on highly specific spellout rules for the honorific pronouns of each language, it becomes *ad hoc* as to which pronoun can be bound in the honorificity-producing configuration in (40b). There is no mechanism or explanation given to exclude unattested proxies of politeness, such as recruiting first/second person for polite address or polite reference[9]. As a result, PPZ do not explain the robust typological trends in recruitment.

Lastly, neither A&N nor PPZ account for multiple-tiered systems of politeness. To do so, the necessary adjustments to the analyses would involve additions of rarefied features, resulting in feature systems like [HON]–[HON+] or [`-formal`][`+formal`][`++formal`], which are in principle imaginable. However, there is scant evidence for such morphosyntactic feature systems. They would predict not only the presence of specialized honorific agreement, but also the presence of specialized *gradable* honorific agreement. To the best of my knowledge, neither type of agreement is typologically attested; the issue of honorific agreement is discussed further in §5.3.3 below. Furthermore, even if there were evidence for such rarefied features, such analyses have no explanatory power over why multiple-tiered systems themselves exhibit recruitment restrictions for each tier (e.g. for singular polite address, why plural is never recruited for a higher tier of politeness than the third person is; the reverse is only ever true).

Hence, whether an analysis of honorific pronouns is based on impoverishment triggered by

[9] [`formal`] is carried over from PPZ's analysis of allocutive markers. Yet, PPZ explicitly acknowledge that allocutivity exhibits important distributional differences from honorific pronouns, so it is unclear why the same morphosyntactic features are used for both phenomena, beyond having social meaning in common. Furthermore, PPZ claim that c^0, the syntactic locus for utterance-level meanings, is covert in languages with honorific pronouns, and overt in languages with utterance-oriented markers of politeness, predicting a two-way typology in this regard. This predicts that there should be *no* language exhibiting *both* allocutivity and a honorificity distinction in pronouns. Yet Basque is precisely such a language: it has both allocutive markers (*-ü-* for formal situations, *-k-/-n-* for colloquial situations) and a familiar/polite distinction in second person pronouns (*hi/zu*).

[HON], or agreement involving [`formal`], a feature dedicated for honorification neither captures—nor explains—significant typological restrictions on the shapes of honorific pronouns. In some cases, the proposed operation may even derive typologically unattested honorification systems. We will see this for A&N's impoverishment mechanism in §5.3.3 below.

5.3.3 An unusual feature, an unusual operation

Even though A&N place restrictions on their impoverishment mechanism, overgeneration is still possible. Here I show that their [HON]-based account predicts unattested patterns regarding i) verbal agreement dedicated to honorification and ii) the existence of certain types of honorification systems. (I do not discuss PPZ in a similar vein, because the authors do not place any restrictions on their agreement mechanism responsible for deriving honorification.)

We start by observing that [HON] is highly unusual as a morphosyntactic feature, having none of the typical properties of its ilk. Let us first consider phonological properties, comparing [HON] to better-established features such as [PL]. In many languages, plurality has dedicated exponence, whether they are pronominal or nominal markers of plurality. A few examples of pronominal plural markers are Yauyos Quechua *-kuna* (Shimelman 2017), Tok Pisin *-pela* (Wurm & Mühlhäusler 1985: 343), Vietnamese *chúng*, Burmese *-tyev* (Cooke 1965); examples of nominal plural markers are Vietnamese *nhung*, Burmese *dowq* (Cooke 1965). However, Chapter 2 showed that this is not the case for [HON]; rather, languages extensively *recruit* existing exponents. Thus, there is little phonological basis for assuming [HON] as a dedicated grammatical feature, if we assume that syntactic features generally have their own exponents.

What about languages which do have dedicated exponents for honorification? In many European languages, the pronoun for honorific address is not recruited from any existing part of the pronominal paradigm. Well-known examples include Spanish *usted*, Portuguese *Você*, Romanian *dumneavoastră*, Dutch/Afrikaans *u*, Polish *Pan/Pani*. For such pronouns, it seems reasonable to assume that [HON] serves to formally distinguish them from their non-honorific counterparts. Indeed, A&N assume dedicated spell-out rules for these honorifics without making use of impoverishment. Let us call such forms *lexicalised honorifics*.

Despite the dedicated exponence, dedicated syntactic reflexes are still lacking. Syntactic features usually drive many formal operations (e.g. movement, concord, agreement). For example, [wh] triggers *wh*-movement. Plural is well-known for having syntactic effects, such as triggering plural concord. Plural number also participates in omnivorous agreement, where one plural marker may cross-reference more than one plural argument (e.g. Nevins 2011), a phenomenon exclusive to singular number.

[HON] does not have any of these syntactic repercussions. As far as I know, no language exhibits syntactic movement triggered by the presence of a honorific pronoun. When a language recruits a pronoun for honorification, the language does not display specialized honorific agreement either[10]. Consider Spanish *usted*: when used for honorific address, it obligatorily triggers 3SG agreement, creating a now-familiar person mismatch (41). For such pronouns, then, A&N must assume that impoverishment selectively targets verbal agreement, but not the pronoun itself.

(41) *¿**Tiene/*tienes** usted la hora?*
have.**3SG/2SG** 2.HON the time
'Do you (POLITE) have the time?'

[10]To explain this, A&N (p. 48) state that impoverishment rules are underspecified for grammatical category, so that they may target verbal agreement as well. But this is tantamount to saying that [HON] does not trigger unique agreement, as it triggers the same process of impoverishment for verbs as it does for pronouns.

However, the lack of dedicated honorific agreement is due to recruitment: many European honorific pronouns originate from nominals referring to purported traits or virtues, such as 'grace', 'lordship', or 'holiness', embodying positive politeness strategies or "approval mechanisms". (Recall that positive face is the desire to be approved of, and positive politeness strategies are acts which reinforce an individual's positive face, such as giving compliments and expressing validation.) Lexical material is recruited and subsequently grammaticalized. Penny (1991: 125) documents this in detail for Spanish, where the expression *Vuestra Merced* 'your mercy' reduced gradually into the present-day pronoun *usted*, used for singular honorific address: *Vuestra Merced* > *vuessa merced* > *vuessarced* > *vuessansted* > *vuessasted* > *voarced* > *vuested* > *usted*.

This process derived pronouns from similar expressions for honorific singular address in related languages, as shown in (42). Note that even though these pronouns were meant for singular address, in many cases the plural possessive pronoun was used, instantiating further instances of plural being recruited for honorification.

(42) Diachronies of lexicalised honorifics

a. Portuguese *Você* (from *Vossemecê* < *Vossa Mercê* 'your.PL mercy');
b. Italian *Lei* (from *La Vostra Signoria* 'your lordship.FEM');
c. Dutch *u* (from *Uwe Edelheid* 'your nobility');
d. Romanian variations on *dumneavoastră* and *dumneata* (from *Dumnia Voastra* 'your.PL grace' and *Dumnia Ta* 'your.SG grace' respectively);
e. Polish variations on *Pan* (from *moj miłościwy Pan* 'my merciful Lord')[11];
f. Hebrew *adoni* (from *adon-i* 'my master').

These honorifics do not trigger unique honorific agreement, but trigger the agreement that accurately reflects their nominal origins. Such mismatches systematically pertain with Portuguese (43a), Italian (43b), Dutch *u* (43c), Romanian *dumneavoastră* (43d), Polish *Pan* (43e), Hebrew *adoni* (43f). In the Italian example (42b), there is an additional mismatch in gender if it is directed towards a male addressee. This is because *Lei* is grammatically feminine, a remnant effect of the possessive pronoun undergoing gender concord with its feminine head noun *Signoria* 'lordship.FEM'.

(43) Lexicalised honorifics recruit existing agreement

a. *Você* ***vem/*vens.***
HON come.PRES.**3SG/2SG**
'You (POLITE) are coming.' (Pedro de Lima, p.c.)

b. *Lei* ***è/*sei*** *malat-**a/*-o.***
HON COP.**3SG/2SG** sick-**FEM/MASC**
'You (POLITE) are sick.' (Pietro Baggio, p.c.)

c. ***Heeft/*hebt*** *u* *een konijn?*
have.**3SG/2SG** HON a rabbit
'Do you (POLITE) have a rabbit?' (Frank Hullekes, p.c.)

d. *Dumneavoastră* ***beţi/*bei*** *vin.*
2.HON drink.**2PL/2SG** wine
'You (POLITE) drink wine.'

e. *Pan chyba* ***zwariował.***
HON perhaps mad.**3**SG.MASC

[11] *Pan* is used for male addressees and *Pani* for female addressees. Furthermore, *Pan* and *Pani* are not "honorific" in the same sense as the others: they are merely alternatives to the familiar forms, used in any situation where the speaker is not intimate with the addressee.

'You (POLITE) must be mad.' (Siewierska 2004: 224)

f. *Im adoni* ***yiten/*titen*** *li 2-3 dowt...*
if 2.HON give.FUT.**3SG.MASC/2SG.MASC** me 2-3 minutes
'If you (POLITE) would give me 2-3 minutes...'

Furthermore, Romanian *dumneavoastră*, being derived from *Dumne* 'lord' and *avoastră* 2PL.POSS, was grammatically 2PL, and so triggers 2PL agreement in (43d). In fact, Romanian has developed a whole host of honorific pronouns which are obviously not dedicated as they vary systematically in person, number, and gender (Stavinschi 2015: 36), in accord with the possessive suffixes which are attached to the feminine noun *dumne* 'lord' via the possessive article *-a*. The agreement of each form reflects their diachronic origins, exemplified for all honorific pronouns as subjects of *a bea* 'to drink' in (44).

(44)	*Dumne-a-ta...*	*bei.* (drink.2SG)	'You (SG, HON) drink.'
	Dumne-a-voastră...	*beţi.* (drink.2PL)	'You (SG/PL, HON) drink.'
	Dumne-a-lui...	*bea.* (drink.3SG)	'He (POLITE) drinks.'
	Dumne-a-ei...	*bea.* (drink.3SG)	'She (POLITE) drinks.'
	Dumne-a-lor...	*beau.* ('drink.3PL)	'They (POLITE) drink.'

In French, a level of politeness higher than *vous* (albeit stilted) is possible with these abstractions, which trigger third person agreement. (45) shows that use of the title *Monsieur* 'sir' for honorific address requires the third person pronoun *il* for co-reference, creating a now-familiar person mismatch. (46) additionally shows that the grammatically feminine abstraction *Votre Altesse* 'your highness' for honorific address towards a man requires the feminine pronoun *elle* for co-reference, creating mismatches in both person and gender.

(45) *Et* ***Monsieur****, qu'est-ce qu'****il*** *désire?*
and Sir what.is-this what.**3SG.MASC** desire
'Sir, what do you (POLITE) desire?'

(46) ***Votre Altesse****, que désire-t-****elle****?*
Your Highness, what desire-t-**3SG.FEM**
'Your Highness, what do you (POLITE) desire?' (Siewierska 2004: 222)

Hence, lexicalised honorifics never trigger dedicated honorific agreement. Rather, agreement simply reflects the original featural specification of the nominal. This also explains why (i) was ungrammatical: agreement with the German honorific *Sie* cannot be second person plural, as *Sie* is third person plural, contrary to what A&N predict.

Thus, [HON] has a very limited range of application: it triggers impoverishment of certain features on certain pronouns and verbs, but not for movement, concord, or agreement, essentially leaving no other morphosyntactic trace behind. In fact, the only operation that [HON] is assumed to trigger—impoverishment—requires some very specific stipulations so that it derives only and all the attested patterns. In particular, LF impoverishment must only target number features, while PF impoverishment must only target person features. I say *must* here, because the two other logical possibilities for impoverishment (that PF impoverishment targets number features, LF impoverishment targets person features) would derive unattested properties of honorification systems. I pursue these possibilities below.

Let us first explore LF impoverishment of person, illustrating with Italian. Recall that Italian uses a third person singular pronoun *Lei* for honorific address. Imagine that [HON] conditions LF impoverishment of the third person feature. In A&N's theory of person, impoverishment of

the third person feature would result in a bare person node, so that the resulting honorific pronoun would be one which is *person*-neutral: such a pronoun can be flexibly used for honorific self-address, honorific address or honorific reference.

However, person-neutral honorific pronouns are unattested. Italian's third person pronoun *Lei* may only be used for honorific address (47a), but not for honorific reference (47b) or honorific self-reference (47c). (Since Italian is a *pro*-drop language, the following cleft forces an overt pronoun to appear.)

(47) *É* ***Lei*** *che mi ispira a studiare.*
COP.3SG HON who 1SG.DAT inspire.PRES to study.INF
a. 'It is you (POLITE) who inspires me to study.'
b. *'It is him (POLITE) who inspires me to study.'
c. *'It is I (POLITE) who inspires me to study.' (Pietro Baggio, p.c.)

Note that the impossibility of (47c) is not due to a mere error in verbal agreement. This is because even in the grammatical sentence (47a), there is already a mismatch between the values reflected in verbal agreement (3SG) and the referential force (2SG).

Honorification systems where a honorific pronoun has flexible person designations are not attested. There are no languages in which a honorific pronoun has available all three interpretations in (47a-c).

Next, consider PF impoverishment of number, illustrating with French (a language which would normally require LF impoverishment of plural). [HON] would condition PF impoverishment of the plural feature, establishing a surface similarity with plural *vous*. However, since LF impoverishment has not applied, this pronoun would *not* be number-neutral, and would only have a plural denotation. The lack of number-neutrality is not the case for French *vous*, as (35) demonstrates (repeated below).

(35) a. *Vous êtes* ***loyal.***
2.HON COP.2PL loyal.**SG**
'You (POLITE) are loyal.'
b. *Vous êtes* ***loyaux.***
2.HON COP.2PL loyal.**PL**
'You all (POLITE) are loyal.'

We do not find any language where a plural pronoun recruited for honorification is used exclusively towards multiple respected addressees. This is not an attested honorification system. Either a language has two unique honorific forms for singular and plural address (as in Spanish *usted/ustedes*), or a recruited form is number-neutral (as in French *vous*, German *Sie*).

In principle, to pair a type of ϕ-feature with a type of impoverishment is a licit analytical move. However, once the "wrong" feature is targeted at the "wrong" level of interpretation, then we derive unattested honorification patterns: namely, honorific pronouns which are person-neutral or exclusively plural.

Chapter 6

Conclusion

At the end of our time together, we assess the results of the long thought experiment presented at the start of this dissertation: What if polite meaning were modeled via recruitment, and not dedicated grammatical features like [HON]?

A happy result is that not much is lost. The OT analysis delivers a factorial typology which mirrors all recruitment asymmetries for all proxands. Depending on which polite grammar is operative, a language may select a number-neutral proxy (plural), a person-neutral proxy (third person singular or inclusive), or a number-neutral *and* person-neutral proxy (third person singular or impersonal). These disparate proxies form a natural class of *semantic defaults*, those semantically underspecified ϕ-featural values which must emerge in contexts involving epistemic uncertainty, arbitrarity, or mixed quantification. The diagnostics for semantic defaulthood emphasize that patterns of morphological exponence cannot be the sole factor informing feature theory—use conditions should be considered as well.

In fact, much more is gained in terms of explanatory power. We are able to explain why semantic defaulthood is a crucial qualifying property for proxies of politeness: using a neutral value allows a speaker to include the respected interlocutor without making direct reference to them, an act which is the linguistic manifestation of negative politeness. The Taboo of Presumptuousness captures the intuition that honorification involves the avoidance of directness, an intuition absent from [HON] analyses. We are also able to explain the shapes of politeness clines: the more presuppositionless a ϕ-featural value is, the more well-suited it is as a negative politeness strategy. The converging shapes of polite proxies and semantic defaults, along with the shapes of politeness clines, may be modeled in a universe with dedicated grammatical features, but cannot be explained.

I would like to take this opportunity to flag two further qualities of the account: it is lean but accommodating. The reader may be reminded of their grandmother's recipe for porridge (or dumplings or beer): simple but infinitely adaptable.

The account's leanness manifests itself in empirical, analytical, and theoretical domains. Empirically, the OT analysis delivers all and only the attested proxies, and excludes all and only the unattested proxies. Deriving such asymmetries with a [HON] account require many additional stipulations, since there are minimal restrictions on the representations that [HON] can sit on. Analytically, ridding our feature inventory of a dedicated grammatical feature for politeness restricts the feature inventory to patterns of exponence which are both detectable and typologically robust, with the welcome consequence that the inventory is not populated with *ad hoc* features. By eschewing [HON], we sidestep unwelcome secondary consequences such as pronoun inventories enlarged by accidental homophony and rampant impoverishment. Rather, this account identifies two independently established mechanisms, semantic defaulthood and negative politeness, as crucial. Theoretically, approximation suggests that the universal inventory of features is economically organized, as features may be "recycled" to serve different functions (e.g. Hale 1986, Biberauer 2018). Learners conservatively postulate the minimal amount of features and make maximal use of them, enabling them to express precisely what they mean even when their language does not provide the relevant grammatical distinctions. This proposal is ambitious because it goes beyond explanatory adequacy: approximation is not only a viable strategy, but also a sensical one. Economy may be an extra-linguistic organizational principle that operationalizes computational efficiency; such so-called "third factors" are not confined to linguistic systems and may be manifest in all subsystems of human cognition, or even in all organic systems (Chomsky 2004, 2005).

Such leanness is a star feature which lends the account easily to many other politeness phenomena. This dissertation has only discussed *pronominal* proxies of politeness, but comparable proxies can be found in the verbal domain as well. First, the same ϕ-featural proxies may crop up in verbal conjugations. Second, approximation pertains beyond the ϕ-featural domain, being

apparent in verbal categories as well. Consider polite commands in Italian and Spanish, which present an enmeshment of these two possibilities. Polite commands are expressed with *third person subjunctive* morphology (in lieu of second person imperative morphology which is used to express non-polite commands; these would be *entra* in Italian (1) or *toma* in Spanish (2)).

(1) *Entri* *pure.*
enter.PRES.**3SG.SUBJUNC** if.you.please
'Please come in.'

(2) *Tome* *su* *libro.*
take.PRES.**3SG.SUBJUNC** your book
'Please take your book.' (adapted from Palmer 2001: 183)

While semantic defaulthood is far from established in this domain, the politeness facts suggest that the subjunctive is the most presuppositionless mood. This is indeed the case that Schlenker (2005) and Siegel (2009) make, *contra e.g.* Quer (2009): that the subjunctive is world-neutral, while the imperative and indicative are world-specific. There is an obvious parallel to be drawn here with ϕ-features; generally, this account predicts that for any set of featural values within the same category, the presupposition-poor value(s) will be recruited for politeness, and the presupposition-rich value(s) will stay far away from politeness.

Third, large pronoun inventories are a politeness phenomenon where several forms variably index formality and/or social setting (potentially along with other factors like sex). These do not exhibit synchronic reflexes of recruitment, but may not be straightforward counterexamples for recruitment. Closer examination may reveal diachronic reflexes of recruitment, which may be in service of either negative politeness (avoidance mechanisms which respect an interlocutor's desire to be unimpeded) or positive politeness (approval mechanisms which respect an interlocutor's desire to be held in esteem). While the latter has not been discussed much in this dissertation, there is no reason to rule it out as an operative pragmatic pressure. A good case study may be the pronominal inventory of Japanese (Christofaki 2018), where none of the forms are obtained from ϕ-featural proxies (although this is unsurprising since Japanese does not morphologically express number distinctions). However, other reflexes of approximation may still be detected. For instance, one of the currently-polite second person pronouns, *anata*, was formed by affixing the distal deictic *a-* to the linker *na-* and *ta* 'side/place', forming a phrase 'that side (far from the speaker)'. Physical distality may serve as a proxy of social distality, hence serving negative politeness pressures. Meanwhile, the once-polite (now informal) second person pronoun, *kimi*, originated from the lexical item 'emperor/lord' (君). As a marker of esteem for the addressee, it would have served positive politeness pressures.

There are surely many more politeness phenomena which piggyback on morphosyntactic resources; the dear reader is reminded that they, too, are potential testing grounds for this account. For now, though, I bid you goodbye!

References

Aalberse, S. (2009). *Inflectional economy and politeness: morphology-internal and morphology-external factors in*, PhD thesis, Universiteit van Amsterdam.

Ackema, P. (2014). Semantic versus syntactic agreement in anaphora: The role of identity avoidance, *Identity Relations in Grammar* pp. 161–196.

Ackema, P. and Neeleman, A. (2018). *Features of Person: From the Inventory of Persons to Their Morphological Realization*, number 78 in *Linguistic Inquiry Monographs*, MIT Press.

Alexiadou, A. (2019). Morphological and semantic markedness revisited: The realization of plurality across languages, *Zeitschrift für Sprachwissenschaft* **38**(1): 123–154.

Alok, D. and Haddican, B. (2022). The formal heterogeneity of allocutivity, *Glossa: a journal of general linguistics* **7**(1).

Arsenijević, B. M. (2014). The agreement of adjectives with the honorific pronoun in Serbo-Croatian, *Philologia Mediana* **6**(6).

Asher, R. E. and Kumari, T. C. (1997). *Malayalam*, Routledge Descriptive Grammars, Routledge, London.

Bale, A., Gagnon, M. and Khanjian, H. (2011). On the relationship between morphological and semantic markedness: The case of plural morphology, *Morphology* **21**: 197–221.

Bashir, E. and Conners, T. J. (2019). *A descriptive grammar of Hindko, Panjabi, and Saraiki*, Vol. 4, Walter de Gruyter GmbH & Co KG.

Baumgartner, R. (1982). *Sociolinguistic Aspects of Persian Pronouns of Address: A Macro Micro Analysis*, PhD thesis, University of Michigan.

Begay, K. R. (2017). *Wailaki grammar*, University of California, Berkeley.

Belay, T. (2015). *A Grammar of Khimt'anga*, PhD thesis, Addis Ababa University, Ethiopia.

Beltrama, A. (2020). Social meaning in semantics and pragmatics, *Language and Linguistics Compass* **14**(9): e12398.

Bentivoglio, P. (2006). Spanish forms of address in the sixteenth century, *Diachronic perspectives on address term systems*, John Benjamins Publishing Company, pp. 177–191.

Besnier, N. (2000). *Tuvaluan: A Polynesian Language of the Central Pacific*, Routledge.

Bhatt, S. K. (2012). Honorifics in Hindi: a morphological, semantic and pragmatical analyses, *Linguistica Brunensia* **60**(1-2): 49–60.

Biberauer, T. (2018). Pro-drop and emergent parameter hierarchies, *in* F. Cognola and J. Casalicchio (eds), *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*, Oxford University Press, pp. 94–140.

Blackwell, S. (2007). Variations in "motherese" pronoun usage, *Studies in Variation, Contacts and Change in English*, Vol. 2, University of Helsinki.

Bobaljik, J. D. (2008). Missing persons: A case study in morphological universals.

Bobaljik, J. D. and Sauerland, U. (2023). About 'us': Clusivity exh, *Proceedings of Sinn und Bedeutung*, Vol. 27, pp. 81–95.

Bobaljik, J. D. and Zocca, C. L. (2011). Gender markedness: the anatomy of a counter-example, *Morphology* **21**(2): 141–166.

Bobuafor, M. et al. (2013). *A grammar of Tafi*, LOT Utrecht.

Bohnhoff, L. E. (2010). *A description of Dii: Phonology, grammar, and discourse.*

Brosig, B. (2018). Pronouns and other terms of address in Khalkha Mongolian, *Philology of the Grasslands*, Brill, pp. 101–111.

Brown, P. and Levinson, S. C. (1978). *Universals in language usage: Politeness phenomena*, Cambridge University Press.

Brown, R., Gilman, A. et al. (1960). *The pronouns of power and solidarity*, Bobbs-Merrill.

Caha, P. (2009). *The nanosyntax of case*, PhD thesis, Universitetet i Tromsø.

Caha, P. (2020). Nanosyntax: Some key features, *Manuscript. Masaryk University, Brno. https://ling.auf.net/lingbuzz/004437* .

Chafe, W. (1990). Uses of the defocusing pronominal prefixes in Caddo, *Anthropological Linguistics* pp. 57–68.

Chatterji, S. K. (1926). *The origin and development of the Bengali language*, Vol. 2, Calcutta University Press.
URL: *https://doi.org/10.4324/9781003481188*

Chierchia, G., Fox, D. and Spector, B. (2012). Scalar implicature as a grammatical phenomenon, *Handbücher zur Sprach-und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science Semantics Volume 3*, de Gruyter.

Chomsky, N. (2004). Biolinguistics and the human capacity, Lecture presented at MTA Budapest on May 17.

Chomsky, N. (2005). Three factors in language design, *Linguistic inquiry* **36**(1): 1–22.

Chomsky, N. and Lasnik, H. (1977). Filters and control, *Linguistic inquiry* **8**(3): 425–504.

Christofaki, R. (2018). *Projected self: the de se across dimensions and beyond pronouns*, PhD thesis, University of Cambridge.

Clyne, M., Kretzenbacher, H.-L., Norrby, C. and Schüpbach, D. (2006). Perceptions of variation and change in German and Swedish address, *Journal of sociolinguistics* **10**(3): 287–319.

Collins, C. and Postal, P. M. (2012). *Imposters: A study of pronominal agreement*, MIT Press.

Comrie, B. (1975). Polite plurals and predicate agreement, *Language* pp. 406–418.

Cooke, J. (1965). *Pronominal Reference in Thai, Burmese, and Vietnamese*, PhD thesis, University of California, Berkeley.

Corbett, G. G. (1983). *Hierarchies, targets and controllers: Agreement patterns in Slavic*, Croom Helm London.

Corbett, G. G. (2004). *Number*, Cambridge University Press.

Corbett, G. G. (2012). *Features*, Cambridge University Press.

Cordona, G. and Suthar, B. (2003). Gujarati, *in* G. Cardona and D. Jain (eds), *The Indo-Aryan languages*, London: Routledge, pp. 658–697.

Croft, W. (2003). *Typology and universals*, Cambridge University Press.

Crowley, T. (1982). *The Paamese language of Vanuatu*, Dept. of Linguistics, Research School of Pacific Studies, The Australian National University.

Cumberland, L. A. (2005). *A grammar of Assiniboine: A Siouan language of the Northern Plains*, Indiana University.

Cysouw, M. (2005). A typology of honorific uses of clusivity, *Clusivity: Typology and case studies of the inclusive–exclusive distinction*, John Benjamins Publishing Company, pp. 213–230.

David, A. B. (2015). *Descriptive grammar of Bangla*, Walter de Gruyter.
URL: *https://doi.org/10.1515/9781614512295*

De Reuse, W. J. and Goode, P. (2006). *A practical grammar of the San Carlos Apache language*, Vol. 51, Lincom Europa.

Deal, A. R. (2015). Interaction and satisfaction in φ-agreement, *Proceedings of NELS*, Vol. 45, pp. 179–192.

Deepak, A. and Baker, M. (2018). On the mechanics (syntax) of indexical shift: Evidence from allocutive agreement in Magahi. Ms., Rutgers University.

Despić, M. (2017). Investigations on mixed agreement: polite plurals, hybrid nouns and coordinate structures, *Morphology* **27**: 253–310.

Dhongde, R. V. and Wali, K. (2009). *Marathi*, John Benjamins Publishing Company.

Dimmendaal, G. J. (1991). The geometry of verb paradigms in teso-turkana, *Paradigms: The economy of inflection*, De Gruyter Mouton.

Donohue, M. (1999). *A Grammar of Tukang Besi*, Mouton.

Duranti, A. and Byarushego, E. R. (1977). On the notion of "direct object", *Haya Grammatical Structure*, Southern California Occasional Papers in Linguistics, chapter 4.

Durie, M. (1985). *A Grammar of Acehnese on the basis of a dialect of North Aceh*, Foris Publications.

Early, R. J. (1994). *A grammar of Lewo, Vanuatu*, PhD thesis, Australian Natl. Univ. Canberra.

Eckardt, R. (2011). Grammaticalization and semantic change.

Eckert, P. (2008). Variation and the indexical field 1, *Journal of sociolinguistics* **12**(4): 453–476.

Elbourne, P. (2001). On the semantics of pronouns and definite articles, *Proceedings of WCCFL*, Vol. 20, Citeseer, pp. 164–177.

Faltz, L. M. (1997). *The Navajo verb: A grammar for students and scholars*, UNM Press.

Firth, R. (1987). *Tikopia-English Dictionary (Taranga Fakatikopia ma Taranga Fakainglisi)*, Sydney University Press for the Anthropological Society of New South Wales.

Frajzyngier, Z. (2012). *A grammar of Wandala*, Vol. 47, Walter de Gruyter.

Givón, T. (2015). *The diachrony of grammar*, John Benjamins Publishing Company.

Goffman, I. (1967). On face-work: An analysis of ritual elements in social interaction, *Interaction Ritual*, Doubleday.

Goswami, G. C. and Tamuli, J. (2003). Asamiya (Assamese), *in* G. Cardona and J. Dhanesh (eds), *The Indo-Aryan languages*, Routledge, chapter 10, pp. 429–484.
URL: *https://doi.org/10.4324/9780203945315-19*

Green, G. M. (1990). The universality of Gricean interpretation, *Annual Meeting of the Berkeley Linguistics Society*, pp. 411–428.

Greenberg, J. (1966). *Language Universals*, The Hague: Mouton.

Greenberg, J. H. et al. (1963). Some universals of grammar with particular reference to the order of meaningful elements, *Universals of language* **2**: 73–113.

Grice, P. (1991). *Studies in the Way of Words*, Harvard University Press.

Gusain, L. (2004). *Marwari*, Lincom Europa.

Gumundsson, H. (1972). *Pronominal Dual in Icelandic*, University of Iceland Publications in Linguistics 2, Reykjavík: Institute of Nordic Linguistics.

Haegeman, L. and Hill, V. (2013). The syntacticization of discourse, *Syntax and its limits* **48**: 370–390.

Hale, K. (1986). Notes on world view and semantic categories: Some Warlpiri examples, *in* P. Muysken and H. van Riemsdijk (eds), *Features and projections*, Foris, Dordrecht, chapter 8, pp. 233–254.

Hale, K. L. (1973). Person marking in Warlbiri, *in* S. R. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*, Holt, Rinehart, and Winston, New York, pp. 308–344.

Hall, D. (2020). The impersonal gets personal: A new pronoun in Multicultural London English, *Natural Language & Linguistic Theory* **38**: 117–150.

Harbour, D. (2014). Paucity, abundance, and the theory of number, *Language* **90**(1): 185–229.

Harbour, D. (2016). *Impossible Persons*, Vol. 74, MIT Press.

Harley, H. (2008). When is a syncretism more than a syncretism? Impoverishment, metasyncretism, and underspecification, *Phi theory: Phi-features across modules and interfaces*, Oxford University Press, pp. 251–294.

Harley, H. and Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis, *Language* **78**(3): 482–526.

Haspelmath, M. (2006). Against markedness (and what to replace it with), *Journal of Linguistics* **42**(1): 25–70.

Haviland, J. B. (1979). Guugu Yimidhirr brother-in-law language, *Language in Society* **8**(2-3): 365–393.

Heim, I. (1991). Artikel und definitheit, *in* A. Stechow and D. Wunderlich (eds), *Semantics: An International Handbook of Contemporary Research*, de Gruyter, Berlin.

Heim, I. (2008). Features on bound pronouns, *in* S. B. David Adger, Daniel Harbour (ed.), *Phi Theory: Phi-Features Across Modules and Interfaces*, Oxford University Press, pp. 35–56.

Heine, B. and Song, K.-A. (2010a). On the genesis of personal pronouns: Some conceptual sources, *Language and cognition* **2**(1): 117–147.

Heine, B. and Song, K.-a. (2010b). Some patterns of grammatical change in personal pronouns, *Language and Cognition* **2**(1): 117–147.

Helm, J. (1965). Bilaterality in the socio-territorial organization of the Arctic Drainage Dene, *Ethnology* **4**(4): 361–385.

Hemmings, C. (2016). *The Kelabit language: Austronesian voice and syntactic typology*, PhD thesis, SOAS University of London.

Hualde, J. I. and De Urbina, J. O. (2003). *A grammar of Basque*, Vol. 26, Walter de Gruyter.

Jaynes, J. (1976). *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, Houghton Mifflin, Mariner Books.

Jensen, J. B. (2012). Forms of address of Brazilian Portuguese: Standard European or Oriental honorifics?, *From Linguistics to Literature*, John Benjamins Publishing Company, pp. 45–66.

Kaur, G. and Yamada, A. (2022). Honorific (mis) matches in allocutive languages with a focus on Japanese, *Glossa: a journal of general linguistics* **44**(1).

Keating, E. and Duranti, A. (2006). Honorific resources for the construction of hierarchy in Samoan and Pohnpeian, *The Journal of the Polynesian Society* **115**(2): 145–172.

Keenan, E. L. (1978). Logic and language, *in* M. Bloomfield and E. Haugen (eds), *Language as a Human Problem*, JSTOR, New York: W.W. Norton and Company, pp. 187–196.

Kilian-Hatz, C. (2008). *A grammar of modern Khwe (Central Khoisan).*

Kitagawa, C. and Lehrer, A. (1990). Impersonal uses of personal pronouns, *Journal of pragmatics* **14**(5): 739–759.

Koul, O. N. and Wali, K. (2006). *Modern Kashmiri Grammar*, Dunwoody Press.

Kratzer, A. (2009). Making a pronoun: Fake indexicals as windows into the properties of pronouns, *Linguistic Inquiry* **40**(2): 187–237.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics, *in* R. Bartsch, J. van Benthem and P. van Emde Boas (eds), *Semantics and contextual expressions*, Foris Publications, Dordrecht, pp. 75–116.

Krishnamurti and Gwynn (1985). *A Grammar of Modern Telegu*, OUP.

Kuiper, A. and Oram, J. (1991). A syntactic sketch of Diuxi-Tilantongo Mixtec, *Studies in the syntax of Mixtecan languages*, Vol. 3, Summer Institute of Linguistics and the University of Texas at Arlington, pp. 179–408.

Laberge, S. and Sankoff, G. (1979). Anything 'you' can do, *in* T. Givón (ed.), *Syntax and semantics 12: Discourse and Syntax*, University of Pennsylvania Press Philadelphia, pp. 419–440.

Lapesa, R. (2000). *Personas gramaticales y tratamientos en español.*

Laughren, M. (2001). What Warlpiri 'avoidance' registers do with grammar, *in* J. Simpson, D. Nash, M. Laughren, P. Austin and B. Alpher (eds), *Forty years on: Ken Hale and Australian languages*, Pacific Linguistics, Canberra, pp. 199–225.

Laughren, M. N., Granites, R. J., Hale, K. L. and Hoogenraad, R. (1996). *A learner's guide to Warlpiri: tape course for beginners. Wangkamirlipa Warlpirilki*, Institute for Aboriginal Development.

Lehman, C. (1977). Imperatives, *Haya Grammatical Structure*, Southern California Occasional Papers in Linguistics, chapter 10.

Lehmann, T. (1993). *A Grammar of Modern Tamil*, Pondicherry Institute of Linguistics and Culture.

Lichtenberk, F. (2003). Inclusive-exclusive in Austronesian: an opposition of unequals, *Clusivity: Typology and case studies of the inclusive–exclusive distinction*, John Benjamins Publishing Company, pp. 261–289.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach, *in* R. B. Scwarze and A. von Stechow (eds), *Meaning, Use, and Interpretation of Language*, Mouton de Gruyter, pp. 303–323.

Listen, P. H. (1999). *The emergence of German polite Sie: Cognitive and sociolinguistic parameters*, Peter Lang.

Lockwood (1975). *The Pronominal Dual* in Icelandic by Helgi Gumundsson, *Journal of Linguistics* **11**(1).

Lohar, G. T. (2020). *A Grammar of Bhojpuri*, PhD thesis, Tribhuvan University.

Lorenzino, G. A. (1998). *The Angolar Creole Portuguese of São Tomé: Its grammar and sociolinguistic history*, City University of New York.

Macaulay, B. K. (2015). The morphosyntax of formality: A typology and inclusion in feature geometry, *University of Pennsylvania Working Papers in Linguistics* **21**(1).

Magier, D. (1983). *Topics in the grammar of Marwari*, PhD thesis.

Malamud, S. A. (2006). *Semantics and pragmatics of arbitrariness*, PhD thesis, University of Pennsylvania Dissertation.

Mayr, C. (2015). Plural definite NPs presuppose multiplicity via embedded exhaustification, *Semantics and linguistic theory*, pp. 204–224.

McCready, E. (2019). *The Semantics and Pragmatics of Honorification: Register and Social Meaning*, Vol. 11 of *Oxford Studies in Semantics and Pragmatics*, Oxford University Press.

McFadden, T. (2017). The morphosyntax of allocutive agreement in Tamil. Ms., Leibniz-Zentrum Allgemeine Sprachwissenschaft.

Melnar (1998). *Caddo verb morphology*, PhD thesis, University of Chicago.

Miller, W. (1980). Speaking for two: Respect speech in the Guarijio of Northwest Mexico, *Berkeley Linguistics Society*, Vol. 6, pp. 196–206.

Mitrović, M. (2025). A syntactic reduction of respect, *Charting honorific and Addressee Morphosyntactic Processes*, number 3.

Miyagawa, S. (2017). *Agreement Beyond Phi*, Vol. 75 of *Linguistic Inquiry Monographs*, MIT Press.

Miyaoka, O. (2012). *A Grammar of Central Alaskan Yupik (CAY)*, Vol. 58, Walter de Gruyter.

Moltmann, F. (2006). Generic one, arbitrary pro, and the first person, *Natural language semantics* **14**(3): 257–281.

Mous, M. (2007). A sketch of Iraqw grammar, *Unpublished manuscript. Leiden University* .

Nevins, A. (2011). Multiple Agree with clitics: Person complementarity vs. omnivorous number, *Natural Language and Linguistic Theory* **29**(4): 939–971.

Noyer, R. R. (1992). *Features, positions and affixes in autonomous morphological structure*, PhD thesis, Massachusetts Institute of Technology.

Osumi, M. (1995). Tinrin grammar, *Oceanic Linguistics Special Publications* (25): i–304.

Oyharçabal, B. (1993). Verb agreement with nonarguments: On allocutive agreement, *in* J. O. de Urbina and J. I. Hualde (eds), *Generative Stuides in Basque Linguistics*, John Benjamins Publishing, pp. 89–114.

Pacheco, C., Carvalho, A. and Scherre, M. P. (2024). When a new pronoun crosses the border: The spread of *a gente* on the Brazilian-Uruguayan frontier, *Languages* **9**(3): 98.

Palmer, F. R. (2001). *Mood and modality*, Cambridge university press.

Pandharipande, R. (2003). Marathi, *in* G. Cardona and D. Jain (eds), *The Indo-Aryan Languages*, Vol. 2, Routledge, chapter 19, pp. 766–802.

Payne, D. L. (1985). *Aspects of the grammar of Yagua: A typological approach*, PhD thesis, UMI Ann Arbor.

Penny, R. J. (1991). *A history of the Spanish language*, Cambridge University Press.

Podobryaev, A. (2017). Three routes to person indexicality, *Natural Language Semantics* **25**(4): 329–354.

Portner, P., Pak, M. and Zanuttini, R. (2019). The speaker-addressee relation at the syntax-semantics interface, *Language* **95**(1): 1–36.

Potts, C. (2005). *The logic of conventional implicatures*, number 7 in *Oxford Studies in Theoretical Linguistics*, Oxford University Press.

Priestly, T. M. S. (1993). Slovene, *in* B. Comrie and G. G. Corbett (eds), *The Slavonic Languages*, Routledge, London, pp. 388–451.

Puškar-Gallien, Z. (2019). Resolving polite conflicts in predicate agreement, *Glossa: a journal of general linguistics* **4**(1): 1–25.

Quer, J. (2009). Twists of mood: The distribution and interpretation of indicative and subjunctive.

Refsing, K. (1986). The Ainu language: the morphology and syntax of the Shizunai dialect.

Rennison, J. R. (1997). *Koromfe: Descriptive Grammars*, Routledge, London and New York.

Rounds, C. (2009). *Hungarian: An essential grammar*, Routledge.

Rubino, C. R. G. (1999). *A reference grammar of Ilocano*, PhD thesis, University of California, Santa Barbara.

Rushforth, S. (1981). Speaking to 'relatives-through-marriage': aspects of communication among the Bear Lake Athapaskans, *Journal of Anthropological Research* **37**(1): 28–45.

Saeli, H. and Miller, C. (2018). Some linguistic indicators of sociocultural formality in persian, *Trends in Iranian and Persian linguistics* **313**: 163.

Samek-Lodovici, V. and Prince, A. (2002). Fundamental properties of harmonic bounding, *Rutgers Center for Cognitive Science, RuCCS-TR-71. Available on Rutgers Optimality Archive, ROA* **785**.

Sauerland, U. (2003). A new semantics for number, *in* R. Young and Y. Zhou (eds), *Proceedings of SALT XIII*, CLC Publications, Cornell, Ithaca, pp. 258–275.

Sauerland, U. (2008). On the semantic markedness of ϕ-features, *in* D. Harbour, D. Adger and S. Béjar (eds), *Phi theory: Phi-features across modules and interfaces*, Oxford University Press, pp. 57–82.
URL: *https://doi.org/10.1093/oso/9780199213764.003.0003*

Sauerland, U. (2009). The presuppositional approach to ϕ-features. lingbuzz/003877.

Sauerland, U., Andersen, J. and Yatsushiro, K. (2005). The plural is semantically unmarked, *in* S. Kepser and M. Reis (eds), *Linguistic evidence*, Mouton de Gruyter, Berlin, pp. 413–434.

Savishinsky, J. S. (1970). Kinship and the expression of values in an Athabascan Bush community, *Western Canadian Journal of Anthropology* **2**(1): 31–59.

Schachter, P. and Otanes, F. T. (1972). *Tagalog Reference Grammar*, University of California Press, Berkeley.

Schiffman, H. F. (1983). *A Reference Grammar of Spoken Kannada*, University of Washington Press, Seattle.

Schlenker, P. (2005). The lazy Frenchman's approach to the subjunctive: Speculations on reference to worlds and semantics defaults in the analysis of mood, *Selected papers from 'Going Romance' 2003, Nijmegen, 20–22 November*, John Benjamins Publishing Company, pp. 269–309.

Senge, C. (2015). *A Grammar of Wanyjirra, a language of Northern Australia*, PhD thesis, The Australian National University.

Shapiro, M. C. (2003). Hindi, *in* G. Cardona and D. Jain (eds), *The Indo-Aryan Languages*, Vol. 2, Routledge, chapter 7, pp. 276–314.

Shelden, D. (1991). Setting a course in galela, *Papers in Papuan Linguistics* p. 147.

Shibatani, M. (1990). *The languages of Japan*, Cambridge University Press.

Shields, J. (1988). A syntactic sketch of Silacayoapan Mixtec, *Studies in the syntax of Mixtecan languages* **1**: 305–449.

Shimelman, A. (2017). *A grammar of Yauyos Quechua*, number 9 in *Studies in Diversity Linguistics*, Language Science Press.

Siegel, L. (2009). Mood selection in Romance and Balkan, *Lingua* **119**(12): 1859–1882.

Siewierska, A. (2004). *Person*, Cambridge University Press.

Sigursson, H. Á. and Egerland, V. (2009). Impersonal null-subjects in Icelandic and elsewhere, *Studia Linguistica* **63**(1): 158–185.

Silverstein, M. (1976). Hierarchy of features and ergativity, *in* R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*, Australian Institutes of Aboriginal Studies, Canberra, pp. 112–171.

Simon, H. (2003). From pragmatics to grammar: Tracing the development of "respect" in the history of the German pronouns of address, *in* I. Taavitsainen and A. H. Jucker (eds), *Diachronic Perspectives on Address Term Systems*, John Benjamins Publishing, pp. 85–134.

Sinha, Y. (2025). *Plurals of politeness and the morphosyntax of number*, PhD thesis, Massachusetts Institute of Technology.

Small, P. (1990). A syntactic sketch of Silacayoapan Mixtec, *Studies in the syntax of Mixtecan languages*, Vol. 2, Summer Institute of linguistics Dallas, chapter 2.

Solnit, D. B. (1997). *Eastern Kayah Li: grammar, texts, glossary*, University of Hawaii Press.

Speas, P. and Tenny, C. (2003). Configurational properties of point of view roles, *Asymmetry in grammar* pp. 315–344.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures., *in* U. Sauerland and P. Statev (eds), *Presuppositions and implicatures in compositional semantics*, Palgrave-Macmillan, Houndmills, pp. 243–281.

Starke, M. (2002). Lectures on nanosyntax. Unpublished Ms., CASTL-University of Tromsø.

Starke, M. (2009). Nanosyntax: A short primer to a new approach to language, *Nordlyd* **36**(1): 1–6.

Stavinschi, A. C. (2015). Romanian, *Manual of Deixis in Romance languages*, De Gruyter, pp. 17–44.

Storch, A. (2014). *A grammar of Luwo*, John Benjamins Publishing Company.

Suharno, I. (1982). *A descriptive study of Javanese*, Research School of Pacific Studies, Australian National University.

Tanangkingsing, M. (2009). *A Functional Reference Grammar of Cebuano*, PhD thesis, National Taiwan University.

Treis, Y. (2008). *A grammar of Kambaata, Part 1: Phonology, Morphology, and Nonverbal Predication*, number 26 in *Cushitic Language Studies*, Rüdiger Köppe Verlag, Cologne.

Valkama, K. (1995). Person marking in Duri, *Studies in Sulawesi linguistics* **4**: 47–95.

van den Berg, R. (1989). *A Grammar of the Muna Language*, Vol. 139 of *Verhandelingen van het Koninklijk Instituut voor Taal-, Landen Volkenkunde*, Foris, Dordrecht.

van den Bergh, J. D. (1953). *Spraakkunst van het Banggais.*

van Driem (1987). *A Grammar of Limbu*, Mouton.

van Klinken (1999). *A grammar of the Fehan dialect of Tetun*, Research School of Pacific and Asian Studies, The Australian National University.

Verheijen, J. A. et al. (1986). *The Sama/Bajau language in the lesser Sunda islands*, Dept. of Linguistics, Research School of Pacific Studies, The Australian ....

Verma, S. (2003). Magahi, *in* G. Cardona and J. Dhanesh (eds), *The Indo-Aryan languages*, Routledge, chapter 13, pp. 547–565.
URL: *https://doi.org/10.4324/9780203945315-22*

Von Fintel, K. and Matthewson, L. (2008). Universals in semantics, *The Linguistic Review* (25): 139–201.

Wakasa, M. (2008). *A descriptive study of the modern Wolaytta language*, PhD thesis, University of Tokyo.

Walker, A. T. (1982). *A grammar of Sawu*, Vol. 13 of *Linguistic Studies In Indonesian and Languages In Indonesia*.

Wang, R. (2023). Honorifics without [HON], *Nat Lang Linguist Theory* **41**: 1287–1347.

Wang, R. (2024). Variable preservation of honorificity after repluralization: A diachronic typology, *Formal Diachronic Semantics*, number 9.

Wang, R. (2025). Tiered honorification in Eastern Indo-Aryan: A [HON]-less proposal, *Semantics and Linguistic Theory*, pp. 1–20.

Wechsler, S. (2011). Mixed agreement, the person feature, and the index/concord distinction, *Natural Language & Linguistic Theory* **29**: 999–1031.

Wechsler, S. and Hahm, H.-J. (2011). Polite plurals and adjective agreement, *Morphology* **21**: 247–281.

Wills, D. D. (1977). Participant deixis in English and baby talk, *in* C. Snow and C. Ferguson (eds), *Talking to Children: Language Input and Acquisition*, Cambridge University Press, chapter 12, pp. 271–298.

Wurm, S. A. and Mühlhäusler, P. (1985). *Handbook of Tok Pisin (New Guinea Pidgin)*, Pacific Linguistics, The Australian National University.

Yadav, R. (1996). *A Reference Grammar of Maithili*, Mouton de Gruyter, Berlin.
URL: *https://doi.org/10.1515/9783110811698*

Yadava, Y. P. (1999). The complexity of Maithili verb agreement, *in* K. M. Rajendra Singh, Probal Dasgupta (ed.), *The Yearbook of South Asian Linguistics*, Vol. 1999, Berlin, New York: De Gruyter Mouton, pp. 139–152.
URL: *https://doi.org/10.1515/9783110245240.139*

Yatsushiro, K., Sauerland, U. and Alexiadou, A. (2017). The unmarkedness of plural: Crosslinguistic data, *in* M. LaMendola and J. Scott (eds), *Proceedings of the 41st annual Boston University Conference on Language Development*, Cascadilla Press, Somerville, MA, pp. 753–765.

Yonezawa, Y. (2021). *The Mysterious Address Term anata 'you' in Japanese*, John Benjamins Publishing Company.

Yousef, S. (2018). *Persian: A comprehensive grammar*, Routledge.

Zilles, A. M. (2005). The development of a new pronoun: The linguistic and social embedding of *a gente* in Brazilian Portuguese, *Language variation and change* **17**(1): 19–53.

Zobel, S. (2014). *Impersonally interpreted personal pronouns*, PhD thesis, Georg-August-Universität Göttingen.

Zobel, S. (2024). Restrictions on the generic interpretation of dedicated impersonal pronouns, *39th West Coast Conference on Formal Linguistics*, Cascadilla Proceedings Project, pp. 695–702.

Zu, V. (2013). Probing for conversation participants: The case of Jingpo, *Proceedings of the Annual Regional Meeting of the Chicago Linguistic Society*, pp. 379–389.

Zweig, E. (2009). Number-neutral bare plurals and the multiplicity implicature, *Linguistics and Philosophy* **32**(4): 353–407.

Zwicky, A. M. (1977). Hierarchies of person, *in* W. A. Beach and S. E. Fox (eds), *Papers from the Thirteenth Regional Meeting of the Chicago Ling. Soc.*, Vol. 13, pp. 714–733.

Zylstra, C. F. (1991). A syntactic sketch of Alacatlatzala Mixtec, *Studies in the syntax of Mixtecan languages*, Vol. 3, Summer Institute of Linguistics and the University of Texas at Arlington.