

MPI Diversity Conference

2 May 2015

A typological overview of strategies that aid in reference tracking

T. Conners, C. Brugman, N. Adams, B. Hefright, A. Gnanadesikan.,
University of Maryland

- People Referring Expressions: holistic account of all grammatical and lexical strategies available to speaker to make reference to a person
- Zero Reference
- Imposter Reference
- Ambiguity reducing strategies



Introduction: typological perspective on theoretical approaches to reference and reference tracking

- Premise 1: no theory of reference should privilege pronouns to the neglect of other expressions that can refer to individuals or co-refer with other expressions that refer to individuals.
- Premise 2: All (language-specific) systems of reference depend on the existence of (language-specific) compensatory systems that aid in reduction of ambiguity. Therefore, understanding systems of reference relies on recognition of these compensatory systems.



Premise 1

- No theory of reference should privilege pronouns to the neglect of other expressions that can refer to individuals or co-refer with other expressions that refer to individuals.
- Expressions such as zeros and “imposters” are not marginal from a typological point of view and should not be marginalized in the theory.
- Kinds of zeros:
 - control zeros (not our focus here)
- Distributional systems of zero types:
 - [Clause-bound] Pro-drop
 - Topic pro-drop
 - Discourse pro-drop



Discourse pro-drop: not a theoretically well reasoned category

- This is a wastebasket category used to dispatch a wide range of zero-distribution types across languages. This is one failure of theory we are seeking to redress.
- Our initial focus is on four languages, all of which fit into this wastebasket, but which differ typologically (and genetically) in important respects.

Chinese

Dhivehi (Maldivian)

Korean

Indonesian



The problem of reference in “discourse pro-drop” languages

- Unlike for languages in which zero is morphologically (“pro-drop”) or syntactically (“topic pro-drop”) conditioned, zero in these languages requires that the referent be recoverable from the discourse (including compensatory strategies) or from the speech setting.
- “Recoverability” requires complex computation from the language, the speech setting, and other cognitive (interpretive) principles.
- Another highly prevalent category of referring expression in these languages, with a similarly complex relationship to referential value, is known as the “imposters” within Indo-European language analysis.

PREs: People referring expressions

- This is a functional category resting crucially on the pragmatic function of referring.
- As far as we know, this is a category that has not been examined *per se* in theoretical, pedagogical, or typological examinations.
- Rather, examination is at the level of (e.g.) pronominal systems or subsystems; pronouns/r-expressions/reflexives; onomastic investigations of proper names; null anaphora.
- Hypothesis: many and diverse linguistic mechanisms are devoted to referring to people that are not devoted to referring to non-people.
- Therefore, expressions referring to non-people will fall out *a fortiori*.

Categories of PREs

- Lexical noun phrases: proper names and descriptive noun phrases
- Personal pronouns
- Pronouns from other languages
- Demonstrative pronouns
- Classifiers, measure terms
- Locative pronouns
- Possessive pronouns
- Reflexive pronouns
- Reciprocal pronouns
- Obligatory agreement expressions
- Zero

Items from these categories may co-occur in some constructions, in some languages.



	Indonesian	Chinese	Korean	Dhivehi
Lexical noun phrases	Yes	Yes	Yes	Yes
Personal pronouns	Yes	Yes	Yes	Yes
Code-switch pronouns	Hokkien, English, Arabic	No	No	Limited English
Demonstratives	Distal/Proximal	No	Yes	3 x distality
Locatives	3 x distality	Restricted	Yes/Noun	Yes, Dem + Loc noun
Possessives	Yes	Yes	No	Genitive pronouns
Reciprocals	Lexical, morphological	Morphological, adverbial	Lexical	Lexical (2)
Reflexives	Multiple types	One free, one bound form	One form	No, repeat referring expression allowed
Null pronouns	Unmarked, S, O, IO	Common S, O, IO	Common S, O, IO	Common, S, O, IO
Imposters	Pervasive	Restricted	Pervasive	Available
Obligatory Agreement	None	None	None	None

Non-CREs: noncanonical referring expressions

- Nulls and imposters have in common that their reference cannot be read off their lexical form. = non-CREs. = “imposters”
- These therefore exist within the functional category of PREs--People-Referring Expressions—along with descriptors, pronouns, etc. PREs include CREs and non-CREs.
- When we explored PREs in these four languages, we found nearly identical inventories but salient dissimilarities in the conditions on their use
- When we explored non-CREs in the four, we found differences in the inventories.

“Imposters”: the gap between lexical and referential values

- Also called *pronoun substitutes* and *open class pronouns*,

An imposter is a notionally X person DP that is grammatically Y person, $X \neq Y$ (Collins and Postal 2012)

Any noun phrase whose denotational value, X, does show exact identity with its referential value, y.

- Mismatch in features such as number, person, gender

Imposters and mismatches

- Collins and Postal introduced this term to cover such cases as (English)
 - ❖ “the royal *we*”: *We* are not amused [said by Queen—mismatch of number]
 - ❖ “the nurses’ *we*”: Have we had our bath today? [said to patient: mismatch of person]
 - ❖ “caregiver speech”: Mommy loves Coco! [said to pet or child: mismatch of person]
 - ❖ Authorial first: In the current work, the authors resolutely disagree.
(mismatch of person)
 - ❖ The Donald Trump: “George Foreman_i. A miracle. A mystery to myself_i.”
(mismatch of person)

Imposters from a typological perspective

- The English inventory of types is common cross-linguistically; however, it is highly contextually restricted from a cross-linguistic point of view
- The T/V distinction (aka the “politeness plural”) in IE is a kind of imposter
- A politeness-based imposter system can encode more distinctions, as in Urdu’s three-way politeness distinction found in 2nd person pronouns
- Imposter systems find their fullest articulation in languages such as Indonesian and Korean, where the inventory of imposters can include



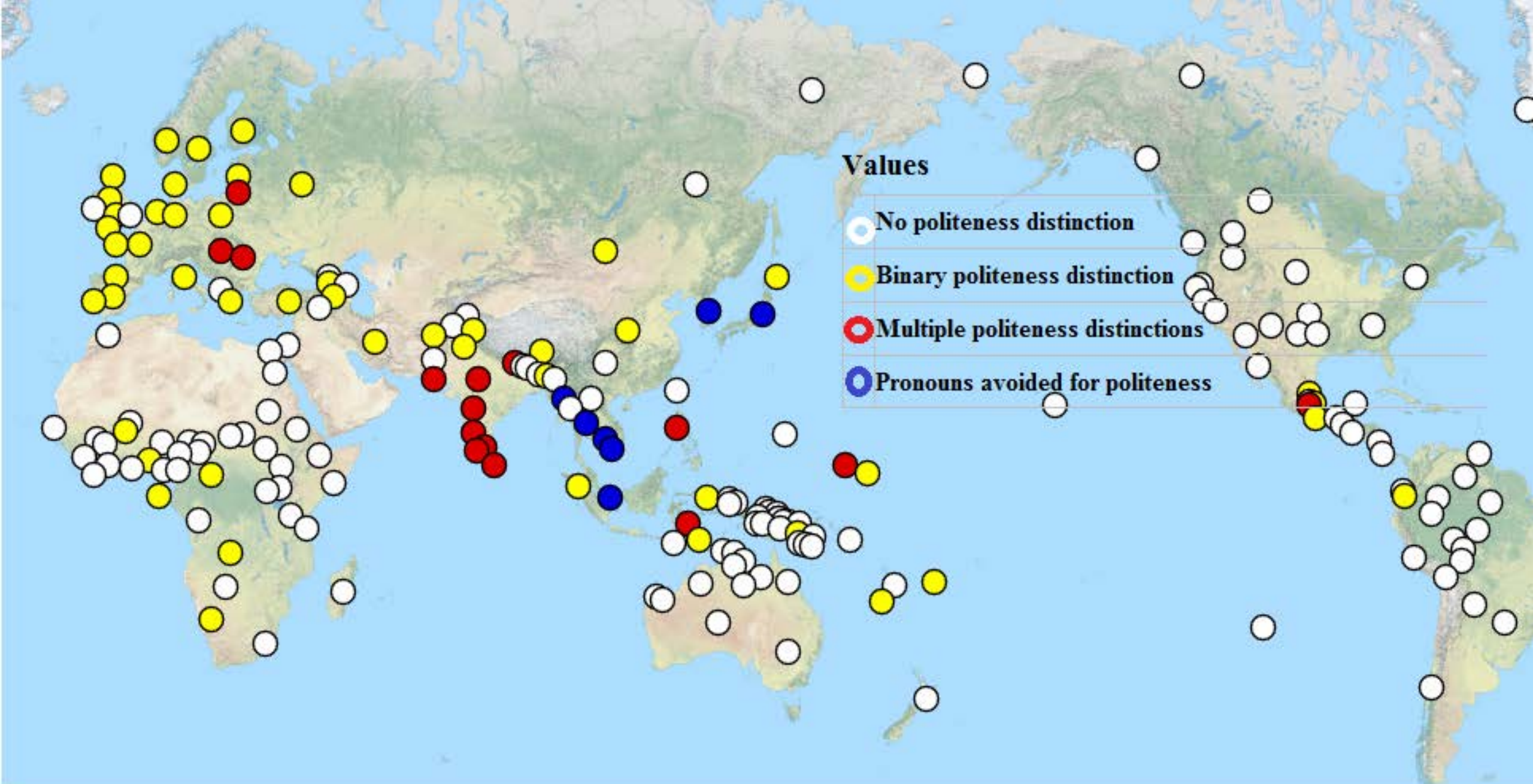
Inventory of imposters for the four languages

- Kinship terms (not restricted to “motherese” register)
- Proper names
- Titles
- Demonstratives (not restricted to presentational constructions)
- Locative adverbs and nouns (don’t have to be deictically grounded, either denotationally or referentially)
- Personal pronouns
- Referring expressions/descriptors



WALS: Pronoun Avoidance

A strategy of pronoun usage which has an effect on the overall shape of the pronominal paradigm. Languages of East and Southeast Asia such as *Japanese*, *Burmese* and *Thai* have a strong sensitivity to politeness in language usage and within their grammars. Speakers have to account for a variety of social distinctions linguistically. Social distinctions between speaker and hearer may reflect relative age, kinship, social ranking, intimacy, and other social features.





○ *Kinship terms*

RAM-200305 [0:13:54]

Nanti kamu dimarahin **Bapaq** lo.

‘I will be angry with you!’

○ *Titles*

BTW-010307 [0:59:52]

Pak ya makan olahan sini mah ya?

‘**You** are eating the menu here right?’

○ *Proper Names*

BTW-010307 [0:43:17]

Saya samaq **Adi** ngoyor ke sono ...

‘I swam with **Adi (you)** over there...’



Imposters not in C&P Typology

- ❖ Demonstrative: **This** is my son-in-law. (#This married my daughter.)
(mismatch of animacy)
- ❖ The animate locative: Do we have that kind of expertise? Not **here!**
[pointing to self] (mismatch of animacy)



Indonesian Demonstratives

- Demonstratives

Ini udah mau bobo!
this already want sleep
'I want to go to bed!'

Tuh mau pergi kan?
that want go PRTCL
'You want to go right.'

Ini and *Itu* w 2nd person reference

- BTW 010307 [0:00:44-0:01:57]
- a: Silakan duduk **tuh (a to c)**!
- ...
- b: Ini ada temen, péngén jalan-jalan ke kampung kitaq.
- a: o gituh?
- b: iyaq.
- a: dari mana **nih (a > c)**?
- c: Saya dari Atma Jaya.
- b: dé péngin ngobrol-ngobrol pakéq basa Betawi.
- ...
- a: orang bahasaq Betawinya udah kagaq **ini (a > a)**...
- a: karena kitaq di sini ...
- a: banyak orang dari Jakarta... udah kagaq anu ... udah kagaq totok kayak orang Betawi

Indonesian Locatives

- Locative pronouns

Sini dah kasi dech.

here already give PRTCL

'I already gave it!'

Sono ikut ga?

there follow NEG

'Are you coming?'



Korean locative noun

그쪽은 어떻게 생각하세요?

geujjok-eun eotteoke saengkakhaseyo

that side-TOP how think (honorific marked verb with polite address ending)

“What do you think?”

Cross-linguistic differences in imposter inventories and conditions

- Chinese: restricted types, highly restricted and marginal in context
- Dhivehi: restricted set (proper names and kin terms), used pervasively
- Korean: full range of types, used pervasively
- Indonesian: full range of types, used pervasively

Imposters cross-linguistically

Type	Indonesian	Chinese	Korean	Dhivehi
Kin term	Common	Rare	Common	Common
Proper name	Common	Never	Sometimes	Common
Title	Common	Sometimes	Common	Rare
Demonstrative	Sometimes	Never	Sometimes	Never
Locative	Sometimes	Never	Sometimes	Never
Descriptive NP	Sometimes	Sometimes	Sometimes	Never

Mismatch types

	Indonesian	Korean	Chinese
Kin term	Person: 1, 2	Person: 1, 2	Person: 2
Proper name	Person: 1, 2	Person: 2	--
Demonstrative	Animacy; Person marked v. unmarked	Animacy; Person marked v. unmarked	--
Locative	Animacy; Person marked v. unmarked	Animacy; Person marked v. unmarked	--
Title	Person: 1, 2	Person: 1, 2	Person: 1
Descriptive NP	--	--	Person: 1,2; Number
Personal Pronouns	--	--	Person: 1

Mismatches are supported by the speech situation as they generally include reference to individuals in the speech situation, Overwhelmingly SAPS

Compensatory strategies

- This is another functional category comprising a very wide range of linguistic strategies that vary cross-linguistically.
- We discuss the ones we have encountered in the four languages
- The notion of “compensation” indexes the idea that rather than aiding in the establishment of reference or coreference, they reduce the ambiguity or uncertainty space for some element whose reference mismatches its lexical features
- They prompt inferences on the part of the hearer

Inventory of compensatory strategies

- Honorifics
 - E.g. distinction in verb forms will point to honored argument referent, either within the speech situation (SAP) or the world spoken of (non-SAP)
- Speech levels
 - E.g. lexical distinctions work similarly with respect to referential ambiguity, with respect to SAPs rather than the world spoken of
 - Why speech level works where register doesn't: registers index social context, speech levels index social identity
- Proximity to deictic origo, where origo can be 1, 2 (Dhivehi)
- Prosody (or its graphic analogues) (e.g. contrastive stress for disjoint reference)
- Optional agreement strategies, e.g. marking topic to aid in anaphoric interpretation
- Extralinguistic information: pointing, eyegaze



Relation to deictic origo: Dhivehi

Locative: used as identity clue

'It is needed **this side**' > 'I need it'

*'This side is hungry'

Demonstrative: used as identity clue

'Did it **here**' > 'I did it'

*'This did it'

Descriptive NP (restricted to 2nd person)

'the parliament member by you' > 'you' (to an MP)

'this parliament member' > 'the one next to me'

*'me'



Korean Honorifics

- Terms of address coded for social distinctions
- Verbal suffix *-si* –ㅅ| is productive
- Used for 2nd/3rd person reference
- Shin (2011): more and more Korean speakers find politeness rules confusing, and default to including honorific markers as much as possible
- Humilifics exist, but rare and not productive.

Speech Levels

- Complete grammatical systems that reflect the relative status of the SAPs. These systems can encode these social relations in their lexicon, morphology, syntax, and phonology. Honorifics may make up part of a speech level system, but other lexical items may also be encoded for relative social status.



Dhivehi Speech Levels

ހަޔަފުޅު ހަޔަފުޅު ހަޔަފުޅު ހަޔަފުޅު ހަޔަފުޅު ހަޔަފުޅު ހަޔަފުޅު ހަޔަފުޅު

Aļugañđu hama hitaş erī iņgētō guļālāfa mivāhaka dakkālan kuḍakoş

1S.DEFER just heart.DAT rise.PST know-Q.HON Call_put-SUCC DEM1-
speech show_put.INF a.bit

I just wanted, [you] know, to call and talk a bit.



Conclusions and Prospects

- Our goal: to lay the groundwork to explore systematically the ways in which languages refer to people (PREs).
- This will help us explain the non-CRE functions of referring to people, which ultimately will allow us to develop a more comprehensive framework of both reference and coreference (inventories and conditions)
 - Expand the typology of ‘imposter’ uses
- More languages, more strategies!: we want to understand better not just non-CREs but also the ways in which languages cue reduction in the referential search space



Reflexive Binding

1. Hanya bapak^[2]bisa mengerti bapak^[2].
'Only father (you) can understand father.'
2. Hanya bapak^[2]bisa mengerti diri bapak^[2].
'Only father (you) can understand father's self.'
3. Hanya bapak^[2]bisa mengerti diri anda.
'Only father (you) can understand yourself.'
4. *Hanya bapak^[2]bisa mengerti dirinya dia.
'Only father (you) can understand himself.'



What is a pronoun? Diagnostics for Indonesian

- *Grammatical Reduplication
 - Reduplication used to express distributive, iterative, multiplicity of types, etc. not available to true pronouns
 - *saya-saya 'we'
 - *gua-gua 'we'
 - But: bapak-bapak, ibu-ibu
- *Modification by Possessor: not available to true pronouns
 - *aku-mu
 - *mereka-ku
 - But: tante-ku, ibu-ku, Wido-mu,
- Determiner (through demonstrative): available to both pronouns and imposters
 - kita ini
 - mereka itu
 - bibi itu