The representation of gender on local person pronouns

Zorica Puškar-Gallien

puskar@leibniz-zas.de
ZAS Berlin

Leibniz-Zentrum
Allgemeine Sprachwissenschaft

Background

Languages in which local person (i.e. 1st and 2nd person) pronouns show gender distinctions have been overlooked, despite abundant literature investigating gender on pronouns (e.g. Audring 2008).
Aim: (i.) Bring such languages together under a single study. (ii.) Provide a model of the representation of ϕ-features on pronouns; derive two main generalisations on the global distribution of gender on pronouns; account for gender distribution within languages and its morphological realisation.
Previous studies: Siewierska (2013) lists 21 languages with gender on local person. Plank and Schellinger (1997) do not offer a formal proposal. Berg (2024), though encompassing a similar corpus, focuses on a different puzzle and lacks a formal proposal.

Data

SG	3 + 2 + 1	Korana (Khoe-Kwadi), Hadza (Isolate), Ngala (Ndu)
SG	3 + 2	Abkhaz (Northwest Caucasian), Arabic, Hebrew, Berber (Riffian, Tacelhiyt, Kabyle, Aures, Tamazight), Amharic, Aramaic, Musey, Kera, Lele, Hausa, Mupun, Kulere, Fyer, Ron (Bokkos, Daffo-Butura), Zari, Miya, Beja (Afro-Asiatic), //Ani, ||Gana, Buga/Khwe, Nama (Khoe-Kwadi), Manambu (Ndu), Kwoma (Sepic), Tunica (Tunica)
SG	3 + 1	Cocama-Cocamilla (Tupian)
SG	2 + 1	Paez (Isolate)
SG	2	Iraqw, Sha, Burunge, Tuareg Berber (Afro-Asiatic), Moken (Austronesian)
SG	1	
PL	3 + 2 + 1	Spanish, Provençal, Shina, Slovenian (Indo-European), Gaagudju (Gaagudju), Berber (Riffian, Tacelhiyt, Kabyle, Tuareg) (Afro-Asiatic), //Ani, ||Gana, Buga/Khwe, Nama, Korana (Khoe-Kwadi), Hadza (Isolate), Emmi (Western Daly), Monumbo (Bogia)
PL	3 + 2	Berber Tumzabt Berber (Aures, Tamazight), Arabic, Hebrew, Beja (Afro-Asiatic), Tunica (Tunica)
PL	3 + 1	
PL	2 + 1	
PL	2	
PL	1	
DU	3 + 2 + 1	Gaagudju (Gaagudju), Slovenian, Lithuanian, (Indo-European), Dumo (Skou), Djeebbana, Burarra, Nakkara (Maningrida), Ngandi, Nunggubuyu, Anindilyakwa (Gunwinyguan), Murrinpatha (Southern Daly), Touo (Isolate), Bora (Boran), Murui Huitoto (Witotoan), //Ani, ||Gana, Buga/Khwe, Nama, Korana (Khoe-Kwadi),
DU	3 + 2	Kamoro (Nuclear Trans New-Guinea)
DU	3 + 1	Cocama-Cocamilla (Tupian)
DU	2 + 1	
DU	2	
DU	1	Wutung (Skou)

Table 1: 62 languages belonging to 20 families with gender distinctions in local person

Generalisations

Global level:

• Gender distinctions on 1st and 2nd person to the exclusion of 3rd possible in the SG and PL.

• Gaps in the following conditions: only 1st person (in SG and PL), only 2nd person non-SG, only 1st and 2nd person non-SG, $3^{rd}+1^{st}$ person (in the PL).

(1) *Generalisation I*:
Having gender distinctions in 1st person singular entails having gender distinctions in the 2nd and/or 3rd person as well.

(2) *Generalisation II*:
Having gender distinctions in 1st person non-singular entails having gender distinctions on 2nd and 3rd peson as well.

$\Rightarrow$ Formal description of 2nd and 3rd person is included in the formal description of the 1st person (Harley and Ritter 2002, Ackema and Neelemann 2013, 2018).

Individual-language level:

• gender distinctions on local person only in the SG

• only non-SG (c.f. Greenberg's 1963 *Universal 37*: A language never has more gender categories in non-singular numbers than in the singular, and *Universal 45*: If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also.)

  – both PL and DU, only DU, only PL (language lacks DU), only PL (language has DU)

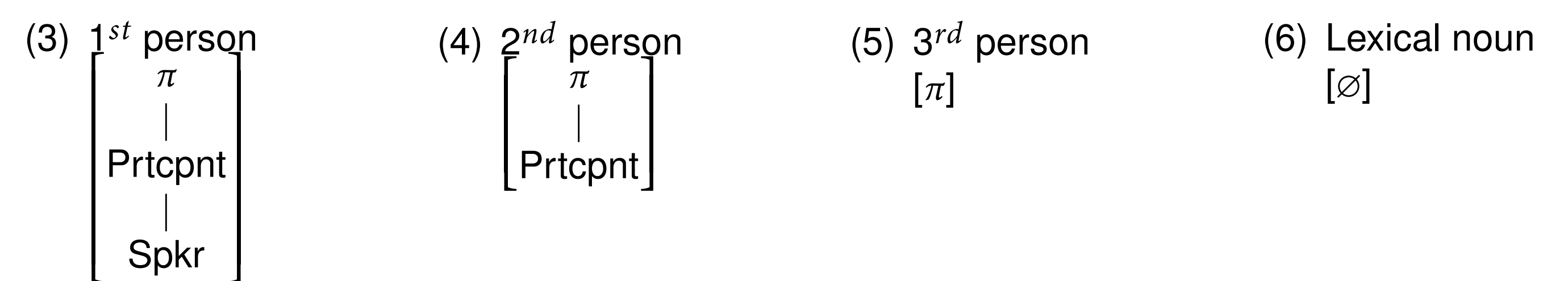• both in the SG and in the non-SG

Morphological level:

• Languages which have been argued to have separate gender morphemes: Spanish, Provençal, Lithuanian, Slovenian, Berber, Amharic, Arabic, Beja, //Ani, ||Gana, Buga, Nama, Korana, Hadza, Dumo, Kwoma, Kamoro, Djeebbana, Burarra, Murrinhpatha, Nunggubuyu, Anindilyakwa, Gaagudju, Touo, Bora, Paez, Tunica, Ngandi, Nakkara, Wutung, Emmi

• Pronouns not segmentable: Abkhaz, Musey, Kera, Lele, Hausa, Mupun, Monumbo, Shina

Selected references

Ackema, P. & A. Neeleman (2018): *Features of Person*. MIT Press. • Adamson, L. J. and E. Anagnostopoulou (2025): 'Gender features and coordination resolution in Greek and three-gendered languages', *LI*. • Béjar, S. and M. Řezáč (2009): 'Cyclic Agree', *LI* 40, 35–73. • Berg, T. (2024): 'Gender, number and person: a three-way interaction', *Linguistic Typology* 28(3), 537–565. • Caha, P. (2021): 'Modeling declensions without declension features. The case of Russian', *Acta Linguistica Academica* 68(4), 385–425. • Deal, A. R. (2015): Interaction and satisfaction in φ-agreement. In: *Proceedings of NELS 45*. 179–192. • Georgi, D. (2013): 'A relativized probing approach to person encoding in local scenarios', *Linguistic Variation* 12(2), 153–210. • Greenberg, J. H. (1963): Some universals of grammar with particular reference to the order of meaningful elements. In: *Universals of Language*. MIT Press, pp. 73–113. • Hammerly, C. (2018): Limiting gender. In: *Gender and Noun Classification*. OUP • Harbour, D. (2016): *Impossible persons*. MIT Press. • Harley, H. and E. Ritter (2002): 'Person and number in pronouns: A feature-geometric analysis', *Language* 78(3), 482–526. • Kratzer, A. (2007): On the plurality of verbs. In: *Event structures in linguistic form and interpretation*. Mouton de Gruyter, pp. 269–299. • Moskal, B. (2015): 'Limits on allomorphy: A case study in nominal suppletion', *LI* 46(2), 363–375. • Nevins, A. (2007): 'The representation of third person and its consequences for the person-case constraint', *NLLT* 25(3), 273–313. • Pesetsky, David (2013): Russian case morphology and the syntactic categories. MIT Press. • Plank, F. and W. Schellinger (1997): 'The Uneven Distribution of Genders over Numbers: Greenberg Nos. 37 and 45', *Linguistic Typology* 1, 53–101. • Puškar-Gallien, Z. (to appear): 'Disassembling and Reassembling Pronouns: A case study of Bosnian/Croatian/Montenegrin/Serbian', *Journal of Slavic Linguistics*. • Siewierska, A. (2013): Gender Distinctions in Independent Personal Pronouns. In: The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology • Smith, P. W., B. Moskal, T. Xu, J. Kang and J. D. Bobaljik (2019): Case and Number Suppletion in Pronouns. *NLLT* 37, 1029–1101. • Trommer, J. (2002): The interaction of morphology and syntax in affix order. In: Yearbook of Morphology 2002. Kluwer, pp. 283–324.

Assumptions

• **A1:** The complexity of the representation of **person** features increases from the 3rd (5) towards the 1st person (3) (Georgi 2013, Nevins 2007, Béjar and Řezáč 2009, Deal 2015).

(3) 1st person
$$\begin{bmatrix} \pi \\ | \\ Prtcpnt \\ | \\ Spkr \end{bmatrix}$$

(4) 2nd person
$$\begin{bmatrix} \pi \\ | \\ Prtcpnt \end{bmatrix}$$

(5) 3rd person
$[\pi]$

(6) Lexical noun
$[\varnothing]$

• **A2:** **Number** is represented by a general number node [#] and a [PL] node below it, as in (7) (Harley and Ritter 2002). Feature [MINimal] may be added to represent dual number. Singular is the absence of number $\Rightarrow$ absence of #P (Kratzer 2007, Nevins 2011, Pesetsky 2013).
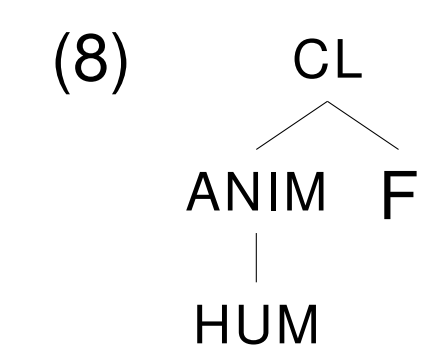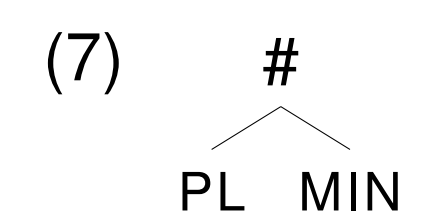
• **A3:** **Gender** includes a general gender node CLASS, a marked feminine value [F] and an animacy [ANIM] and humanness [HUM] specification (8) (see also Hammerly 2018, Caha 2021, Adamson and Anagnostopoulou 2025, Puškar-Gallien to appear).

(7)
```
      #
     / \
   PL   MIN
```

(8)
```
       CL
      /  \
   ANIM   F
     |
    HUM
```

• **A4:** Each feature type projects a phrase. No root – n is the base (Moskal 2015). Person is lower than number (Noyer 1992, Trommer 2002, Harbour 2016, Ackema and Neelemann 2018, van Urk 2019). Gender as CLP projected above #P (Puškar-Gallien to appear).
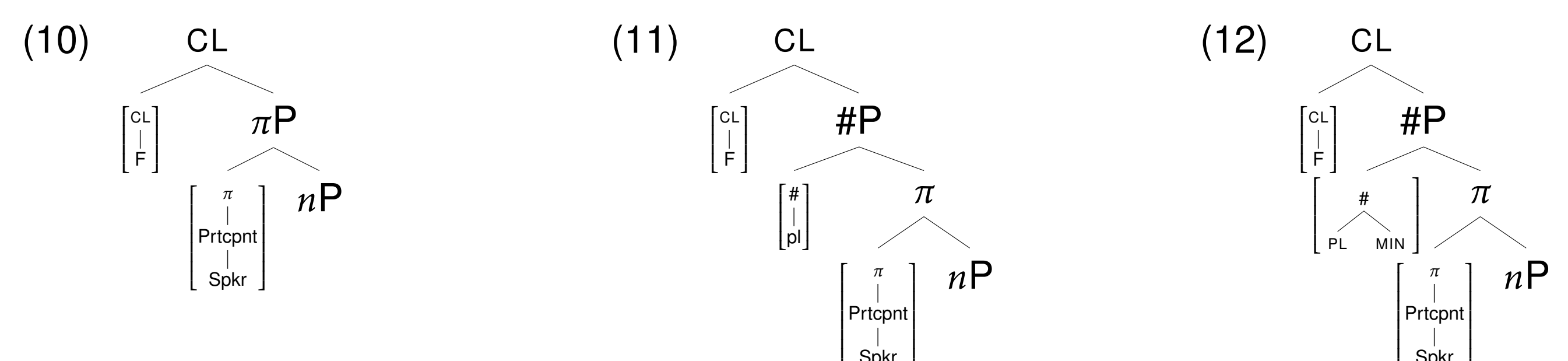
(9) $[_{KP} K [_{CLP} CL [_{\#P} \# [_{\pi P} \pi [_{nP} n]]]]]$

• **A5:** Morphology interprets syntactic structures (Distributed Morphology, c.f. Halle and Marantz 1993), Vocabulary Insertion dischages syntactic features root-outwards; **cyclic realisation** of the phrases in (9) (Bobaljik 2000).

• **A6:** #P is a cyclic domain in the sense of Moskal (2015), Smith et al. (2019).

Analysis

The internal structure of 1st person: singular (10), plural (11) and dual (12)

(10)
```
      CL
     /  \
  [CL    πP
   |    /  \
   F]  π    nP
       |
     Prtcpnt
       |
      Spkr
```

(11)
```
      CL
     /  \
  [CL    #P
   |    /  \
   F]  #    π
      [pl]  |  \
           Prtcpnt  nP
            |
           Spkr
```

(12)
```
      CL
     /  \
  [CL    #P
   |    /  \
   F]  #    π
      / \   | \
    PL  MIN π  nP
          Prtcpnt
            |
           Spkr
```

Global Consequences:
Generalisation I: If gender is present on 1st person, it will be present on other persons, c.f. A1.
Generalisation II: Gender does not combine freely with any person in the non-singular number, due to cyclicity and since #P introduces a locality barrier (c.f. A2, A3, A4, A5).

Language-internal consequences: Having gender present in some environments justifies postulating it across the board (e.g. languages with gender distinctions only on 2nd person SG distinguish between 3rd person M and F *object* pronouns: Iraqw, Burunge (Cushitic), Sha (Chadic), c.f. Mous 1993, Kiessling 1994, Jungraithmayr 1970). The absence of gender in certain environments depends on language-internal factors (markedness toleration, c.f. Weisser in press, principle of informativity, c.f. Berg 2024, diachrony, c.f. Plank and Schellinger 1997, etc.)

Morphological consequences: A4 and A5 predict that π should be realised closest to the nominal base, followed by # and then by gender (*Mirror Principle*, Baker 1985). Languages in Table 2 confirm this prediction. Others: Spanish, Provençal, Lithuanian, Slovenian, (most) Berber languages, Arabic, //Ani, Buga, Nama, ||Gana, Dumo, Djeebbana, Murrinpatha, Nunggubuyu, Anindilyakwa, Touo, Bora, Paez, Nakkara, Emmi. The rest are under scrutiny.

SG	Korana	Hadza	Beja	Gaagudju
1			(un)àni	ngaanj
1M	ti-re	ʼono		
1F	ti-ta	ʼono-ko		
2				ngiinja
2M	sa-ts	te	(um)baʼu:k	
2F	sa-s	te-ko	(um)baʼu:k	
3M	ll'di-b	bami	(um)baʼu:	ngaayu
3F	ll'di-s	bo-ko	(um)baʼu:	naawu
3C	ll'di-ʼi			
DU				
1INM	sa-kham			manaa-mana
1INF	sa-sam			manaa-njdja
1INC	sa-m			
1EXM	si-kham			ngaa-mana
1EXF	si-sam			ngaa-njdja
1EXC	s-im			
2M	sa-kharo			nginjaa-mana
2F	sa-saro			nginjaa-njdja
2C	sa-khaoo			
3M	ll'di-khara			nowoo-mana
3F	ll'di-sara			nowoo-njdja
3C	ll'di-kha			
PL				
1			(an)hìnin	
1INM	sa-tje	ʼuni-bʼi		man-aada
1INF	sa-se	ʼone-beʼe		man-eemba
1INC	sa-da			
1EXM	si-tje	ʼu-bʼi		ng-aada
1EXF	si-se	ʼo-beʼe		ng-aamba
1EXC	si-da			
2M	sa-kao	ʼiti-bi	(am)baʼa:k(na)	nginj-aada
2F	sa-sao	ʼete-be	(am)baʼa:k(na)	nginj-eemba
2C	sa-du			
3M	ll'dku	bʼi	(am)baʼa:	now-ooda
3F	ll'dide	beʼe	(am)baʼa:	now-oomba
3C	ll'dine			

Table 2: Personal pronouns in Korana (Khoe-Kwadi, South Africa; Siewierska 2013), Hadza (isolate, Tanzania; Sands (2013: 270)), Beja (Cushitic, Sudan; Vanhove 2014: 16), Gaagudju (Gaagudju, Australia; Harvey 2002)