PERSON AND GENDER IN PRONOUN PARADIGMS: A SEMANTIC ACCOUNT OF A MORPHOLOGICAL PATTERN*

Bronwyn M. Bjorkman, Isabelle Boyer, Elizabeth Cowper, Daniel Currie Hall, Louise Koren & Dan Siddiqi

Queen's University (Bjorkman); Carleton University (Boyer, Koren, Siddiqi); University of Toronto (Cowper); Saint Mary's University (Hall)

1. Introduction

- Pronominal gender marking (Greenberg 1966; Siewierska 2013)
 - often only in 3rd person
 - often only in 3rd person singular
- Some non-explanations
 - Feature geometry (Harley and Ritter 2002): Perhaps gender features are dependents of whatever characterizes nonparticipants?
 BUT 3rd person is underspecified; no feature on which Gender features could depend. Also, not all languages restrict gender to 3rd person.
 - Impoverishment: Perhaps a rule deletes gender features in the context of [PARTICIPANT]? Works well as a stipulative description, but no explanation for the asymmetry.
- A possible approach: Syntactic structure of participant vs. non-participant pronouns
 - Perhaps gender features are high in the nominal spine, on a head that participant pronouns don't have?
 - But what about languages that mark gender on participant pronouns?
- Our proposal: The syntactic account is part of the answer, but the asymmetry arises from a difference in the semantic type of different person features.
 - Participant ϕ Ps are of type e; nonparticipant ϕ Ps are of type $\langle e,t \rangle$
 - Gender features are of type (et, et), so compose only with predicates.

^{*}The research presented here is financially supported by the Social Sciences and Humanities Research Council of Canada. The ordering of the authors' names is merely alphabetical. We continue to be grateful to Betsy Ritter and Heather Bliss for making their database available to us.

2. The phenomenon

2.1 Previous typological work on gender, number, and person

- Siewierska (2013)
 - "Gender oppositions in personal pronouns are characteristic of the third rather than the first or second person."
- Greenberg (1966)
 - Universal 44: If a language has gender distinctions in the first person, it always has gender distinctions in the second or third person, or in both (Greenberg 1966: 76).
 - Universal 37: A language never has more gender categories in non-singular numbers than in the singular (Greenberg 1966: 75).
 - Universal 45. If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also (Greenberg 1966: 76).
- Plank and Schellinger (1997: 94)
 - A gender distinction in the singular authorizes the same gender distinction in non-singular numbers of the corresponding person.
 - A gender distinction in 3rd person singular authorizes the same gender distinction in 1st and/or 2nd person non-singular.
 - A gender distinction in 3rd person non-singular needs no special authorization.
 - If gender is limited to non-singular, a gender distinction in 3rd person authorizes the same gender distinction in 2nd person, which in turn authorizes the same gender distinction in 1st person.

2.2 Confirmation in a new survey

- Beginning with Bliss and Ritter's (2009) pronouns database, with a few additions
- Currently extends to 112 languages
- Data drawn from published grammars, sometimes supplemented by articles
- The specific question for today: If there is a syncretism in a pronominal paradigm, is it more likely that a distinction in person will be preserved at the expense of a distinction in gender, or vice versa?
- What counts as gender marking in a pronominal system? For the present purposes, any distinction that could not be characterized in terms of either
 - participant features (e.g., inclusive vs. exclusive)
 - politeness/honorification

• What semantic distinctions did we find? For the most part, gender systems included either masculine/feminine or animate/inanimate (? or human/nonhuman??).

The numbers

- 112 languages in the survey; 55 mark gender on at least some pronouns.
- 42 of 55 mark gender **only** in non-participant (i.e., 3rd person) pronouns.
- 17 of 42 mark gender only in the singular (in boldface).
- Only **one** language marks gender on participants but not 3rd persons
- Of the remaining 12:
 - 6 mark gender in all persons
 - 6 mark it in 2nd and 3rd person but not in 1st person.

PARTICIPANTS ONLY	PARTICIPANTS A	ND NON-PARTICIPANTS	NON-PARTICIPANTS ONLY
2nd only	1st, 2nd & 3rd	2nd & 3rd only	3rd only
Iraqw	Djingli, Ngandi, Rikbaktsa, Slovenian (Lithuanian) (Spanish)	Arabic, Bandjalang, Hausa, Hebrew, Tamazight, Tunica	Ainu, Albanian, Arapesh, Asheninca, Awtuw, Basque, Catalan, Chinook, Cubeo, Czech, Dieri, Dutch, (Old) English, German, Godie, Greek, Halkomelem, Hinuq, Ho, IsiXhosa, Kaingang, Kannada, Latin, Latvian, Lushootseed, Marshallese, Mixteco, Mohawk, Pakaasnovos, Palauan, Polish, Pomo, Romanian, Somali, Sotho, Swedish, Telugu, Welsh, Wolaytta, Xokleng, Yimas, Zapoteco
1	6	6	42

Table 1: *Gender marking by person*

(1) Czech pronouns: Gender marked in 3rd person only (Janda and Townsend 2002)

		SING.	PLUR.
1		já	my
2		ty	vy
	MASC.	on	oni
3 <	FEM.	ona	ony
	NEUT.	ono	ona

(2) Awtuw pronouns: Gender marked in 3rd-person singular only (Feldman 1986)

	SING.	DUAL	PLUR.
1	wan	nan	nom
2	jen	an	om
$2 \\ 3 \begin{cases} \text{FEM.} \\ \text{NON-FEM.} \end{cases}$	tej rej	ræw	rom

3. The theoretical challenge

Is gender just less 'useful' in the 1st and 2nd person than in 3rd?

- Author and Addressee are uniquely identifiable by their roles in the discourse.
- Gender can help to disambiguate the reference of non-participant pronouns.

BUT:

- This is not an explanation. It makes no testable predictions.
- Gender marking can do more than resolve reference ambiguities
- (3) Gender and register in Japanese first-person pronouns (Miyazaki 2004: 257, citing Ide 1997: 73)

	Men's speech	Women's speech
Formal	watakushi watashi	watakushi atakushi
Plain	boku	watashi atashi
Deprecatory	ore	Ø

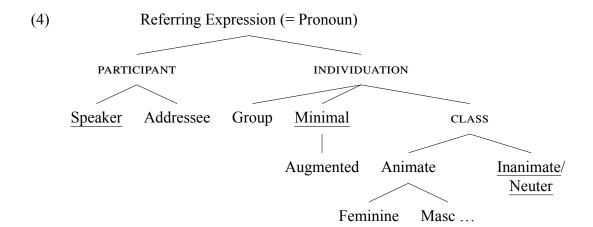
3.1 Formal assumptions

- General syntactic framework of Minimalism (Chomsky 1995 et seq.)
- Distributed Morphology (Halle and Marantz 1993, 1994; Harley and Noyer 1999):
 - Late insertion: VIs are inserted post-syntactically.
 - Underspecification: A given VI may be compatible with several paradigm slots.
 - Competition: The most highly specified compatible VI is selected.
- Neo-parametric approach to formal features and the syntactic structures they project (Cowper and Hall 2014, 2017):
 - Features are not universal.
 - The mechanism by which features are acquired is universal.
 - The syntactic configuration of features is not universal, but is constrained by the requirements of feature-checking and semantic composition.

3.2 Two possible but unsuccessful accounts

3.2.1 Feature geometry

- Do the interactions between gender and person follow from dependency relations among features? A feature-geometric account of the interaction of Gender with Number (Harley and Ritter (2002: 486)).
 - Greenberg's (1966: 74) Universal 36: "If a language has the category of gender, it always has the category of number." →
 - class (gender) is geometrically dependent on Individuation (number)



BUT this won't work for person, especially if the features are privative:

- Third person is unmarked relative to 1st and 2nd person. It lacks [PARTICIPANT] (Harley and Ritter 2002), [PERSON] (Bonet 1991), or [PROX] (Ackema and Neeleman 2013).
- There is no feature that third-person pronouns have that participant pronouns lack.
- So there is no feature that third-person-only Gender could be dependent on.

If the features are binary (e.g., Noyer 1992; Harbour 2016; Cowper and Hall 2019), maybe gender could be a dependent of [-participant]?

- Maybe so, but then why aren't there just as many languages where gender is a dependent of [+participant]?
- It's more common that marked features have dependents. If anything, we'd expect that gender would be more commonly restricted to [+participant] pronouns.

AND, remember that not all languages restrict gender to third persons. So the interaction of gender and person features is not crosslinguistically consistent.

- If the feature geometry is universal, then gender can't be a dependent of 3rd person. Some languages mark gender on participants.
- If the feature geometry is language-specific, then it's completely mysterious why the distribution of patterns is as asymmetric as shown in Table 1.

Moving on ...

3.2.2 Impoverishment

- In Distributed Morphology, the operation of **Impoverishment** (Bonet 1991; Halle and Marantz 1993) can delete features.
- This makes it possible to stipulate that some marked feature (like [GENDER]) is realized only if another marked feature (like [PARTICIPANT]) is **absent**.
- Aha! Just what we were looking for! So, a language-specific impoverishment rule can delete gender features in the context of [PARTICIPANT].
- (5) Korana (Central Khoisan): No impoverishment, gender marked in all persons

	MASC.	FEM.
1st	tire	tita
2nd	sats	sas
3rd	ll'dib	ll'dis

(6) Québécois: [FEM] $\rightarrow \emptyset$ / [PART], gender marked only in 3rd person

	MASC.	FEM.
1st	3	3
2nd	t	t
3rd	i	a

- But wait a minute. What if an impoverishment rule deleted [PARTICIPANT] in the context of a marked gender feature, like [FEMININE]?
- (7) Unattested: [PART] $\rightarrow \emptyset$ / [FEM], person marked only in non-feminine.

	MASC.	FEM.
1st	3	a
2nd	t	a
3rd	i	a

- This impoverishment rule is as simple and natural as the one deriving the Québécois pattern in (6), but the pattern in (7) is unattested in any of the languages we have looked at.
- So, not what we were looking for after all.

4. Proposal: A syntactic account, along with independently motived constraints on semantic composition

- In many languages, third-person pronouns are, or pattern with, demonstratives (Ritter 1995).
- Perhaps in some languages, third-person pronouns have a layer of structure—perhaps D—that participant pronouns lack.
- Perhaps gender features are on that head:



- But what prevents the D head from combining with a ϕ P that contains [PARTICIPANT]?
- And if a D head could select a ϕ P without [PARTICIPANT], why couldn't a D head in a different language select a ϕ P with [PARTICIPANT]?

A possible motivation:

- What if participant and non-participant ϕ Ps had different semantic types? Then it would be unsurprising if they had different potential for composing with further syntactic heads.
- In fact, this has been proposed for Heiltsuk (Bjorkman et al. 2019), and for Marshallese (Cowper and Hall 2022).

4.1 The Heiltsuk story

(9) Heiltsuk¹ demonstratives (Rath 1981: 87–88, 91, cited in Bjorkman et al. 2019)

CLITIC FULL GLOSS

I
$$ga\chi^w$$
 $g\acute{a}q^w$ 'this' (here with me)

II $gats\chi^w$ $g\acute{a}tsq^w$ 'this' (here with me, invisible)

III $qu\chi^w$ $q\acute{u}q^w$ 'that' (there with you)

IV $qu\chi^w ts\chi^w$ $q\acute{u}\chi^w tsq^w$ 'that' (there with you, invisible)

V $qi\chi^w$ $q\acute{u}q^w$ 'that' (over there / under discussion)

VI $qits\chi^w$ $q\acute{u}tq^w$ 'that' (over there / under discussion, invisible)

VII $qki\chi^w$ $qk\acute{u}q^w$ 'that' (absent / gone)

- Ontology of persons: $\{i, u, o, o', o'', o''', \dots\}$ (Harbour 2016)
- Grammatical person features: [\pm author, participant] (Cowper and Hall 2019, and a locative element χ , adapted from Harbour 2016)

(10)
$$[\![\chi]\!] = \lambda x \cdot \lambda y \cdot \text{NEAR}(y, x)$$

(11) Participant π heads are of type e, and can combine with χ .

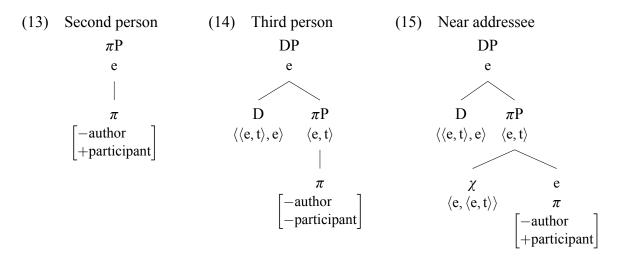
a.
$$\begin{bmatrix} +author \\ -participant \end{bmatrix} = i$$
 b. $\begin{bmatrix} -author \\ +participant \end{bmatrix} = u$ c. $\begin{bmatrix} +author \\ +participant \end{bmatrix} = iu$

¹Heiltsuk (Wakashan), known as Haíłzaqvla by its speakers, is the language of the Heiltsuk nation, located on the Pacific coast in the province of British Columbia. All data cited here comes from Rath (1981), a descriptive grammar and dictionary.

The 3rd-person π head [-author, -participant] denotes the property of not being a discourse participant, and is of type (e, t)

(12)
$$\begin{bmatrix} -\text{author} \\ -\text{participant} \end{bmatrix} = \lambda x \cdot x \in \{o, o', o'', o''', \ldots\}$$

- A Heiltsuk participant πP can be an argument by itself (13).
- A nonparticipant πP needs a D head to derive an e-type argument (14).
- So does a participant πP that has combined with χ (15)



4.2 Where does gender fit in?

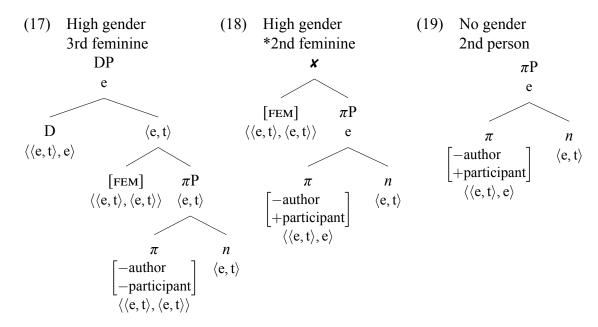
- How do gender features work, semantically?
 - Plausibly, (et, et)
 - Which means that they can compose with 3rd-person— $\langle e, t \rangle$ — πPs , but not with participant, e-type πPs

(16) a.
$$[MASC] = \lambda F.\lambda x.F(x)$$
 & MASCULINE $(x)^2$ b. $[FEM] = \lambda F.\lambda x.F(x)$ & FEMININE (x)

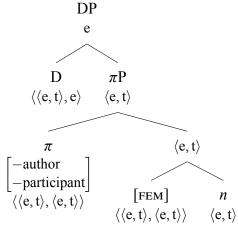
OK, fine, but what about languages that also mark gender in participant pronouns?

²The predicates MASCULINE(x) and FEMININE(x) can be satisfied based on biosocial and/or grammatical gender, depending on the system of a particular language. We assume that other kinds of gender/animacy/noun-class features have the same compositional property. If gender/animacy/noun class should be analyzed in different terms for a given language—as for example animacy in Blackfoot has been analyzed as a form of nominal aspect by Ritter (2014)—it would fall outside the scope of our proposal here.

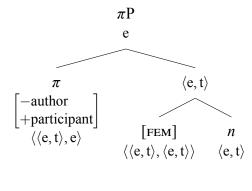
- Suppose that pronouns have a nominal core (n), which is of type (e, t).
- Assume that π heads don't directly denote author, addressee, or others. Instead:
 - π heads with [+author] or [+participant] are of type \langle et, e \rangle .
 - 3rd-person π heads are of type $\langle et, et \rangle$.
- Finally, suppose that gender can be either high (above person), or low (on n)
 - If gender is high, it will combine with 3rd persons, but not with participants.
 - If gender is low, it will combine with *n*, and the result can combine with any set of person features.



(20) a. Low gender 3rd feminine: DP



b. Low gender 2nd feminine:



- This account supports the view that the locus of gender is parametric (Ritter 1993; cf. Kramer 2016).
- Per Ritter (1993), Hebrew full DPs have low gender. Hebrew 2nd-person pronouns also mark gender, as expected if (pro)nominal gender is consistently low.

	MASC.	FEM.
1st	i	i
2nd	xa	ex
3rd	0	a

Table 2: Hebrew singular possessive clitic pronouns

4.3 Interim summary of person and gender in pronouns

- Two positions for gender \rightarrow two kinds of languages with gender in pronouns
 - Languages that can mark gender only in the third person (42 out of 55)
 - Languages that can mark gender in all persons (6 out of 55)
- Unexplained patterns, so far due to accidental syncretism:
 - Seventeen languages with 3rd-person gender only mark it only in the singular.
 - Six languages of 55 mark gender in 2nd and 3rd person, but not in 1st person.
 - One language of 55 marks gender **only** in 2nd person.
 - No languages mark gender in 1st person but not 2nd.

5. Remaining puzzles and future directions

- Why is 2nd person more likely to mark gender than 1st person?
- Why are singular pronouns more likely to mark gender than plurals?

5.1 An asymmetry between participants?

- 13 of 55 languages with pronominal gender mark gender on participants.
- 6 of the 13 mark gender on all participants.
- 7 of the 13 mark gender on 2nd person but not 1st person participants
- None mark gender on 1st person but not on 2nd person.
- Sometimes, the gender marking seems not purely pronominal (Lithuanian, Spanish)

(21) Lithuanian nominative pronouns (Ambrazas 1997: 166, Plank and Schelling
--

		SING.	DUAL	PLUR.
1	MASC. FEM.	àš	mù+ du mù+ dvi	mẽs
2	MASC. FEM.	tù	jù+ du jù+ dvi	jữs
3	MASC. FEM.	jìs jì	jiẽ+du, juõ+du jiẽ+dvi, juõ+dvi	jiẽ jõs

- For us, all 13 languages with participant pronoun gender have low gender.
- Is the absence of languages with gender on 1st but not 2nd person pronouns accidental, or systematic?
- Is there another structural distinction that might explain a systematic gap of this sort?
- More languages, and careful examination of languages with participant pronoun gender, are needed.

5.2 Interaction with number

As noted in section 2, gender also interacts with number. Of the 55 languages in our sample that mark gender on pronouns at all, 17 mark gender only in the singular. Only two, Marshallese and Palauan, mark gender only in the plural. This asymmetry is even more pronounced in the 42 languages where gender contrasts are expressed only in the third person. Of these 42 languages, 23 mark gender in both third person singular and plural, while 17 mark gender only in the third person singular, and two mark it only in the plural.

Recall that in section 3.2.1 we saw that Harley and Ritter's (2002) feature geometry encoded the interaction of gender with number, but not with person. While our account links the interaction of person and gender to semantic interpretation, the interaction of number and gender remains to be explored.

5.3 Future directions

The work reported here is part of a broader research program on what we have called **morphological upstaging**. In Bjorkman et al. (2022) we reported that person is the feature most likely to 'upstage' another in pronoun/demonstrative paradigms, followed by number. Here, we have addressed an apparent tendency to realize person and number at the expense of gender. We will turn in future work to the realization of ϕ -features in verbal agreement, where number so far appears to be the feature most likely to upstage the others.

In the longer term, we hope to identify which upstaging patterns are robust enough to need explanation, and determine whether all upstaging patterns can be attributed to indepen-

ALL NUMBERS	SINGULAR ONLY	PLURAL ONLY
Ainu, Albanian, Arapesh, Asheninca, Basque, Catalan, Czech, Godie, Greek, Halkomelem, Ho, IsiXhosa, Kaingang, Kannada, Latin, Latvian, Mixteco, Mohawk, Pakaasnovos, Polish, Romanian, Sotho, Yimas, Zapoteco	Awtuw, Chinook, Cubeo, Dieri, Dutch, (Old) English, German, Hinuq, Lushootseed, Pomo, Somali, Swedish, Telugu, Welsh, Wolaytta, Xokleng	Marshallese, Palauan
23	16	2

Table 3: Gender marking by number in languages where gender is marked in 3rd person only

dent factors such as syntactic structure, semantic type requirements, or some other as-yet unnoticed factor. It is in principle possible that some upstaging patterns may have to be attributed to irreducible hierarchies among features—hierarchies that stipulate that certain features must be realized even at the cost of others. Such hierarchies could easily be encoded as extrinsic ranking between equally specific Vocabulary Insertion rules, but for us, such an account would be the analysis of last resort.

The semantically based structural account of the interaction of person and gender features in pronouns given above shows one way to derive a 'preference' for one kind of feature over another without stipulating Impoverishment rules or feature hierarchies, and without changing the basic assumptions of vocabulary insertion in Distributed Morphology. To the extent that it is correct, it supports the neoparametric view of formal features and their mapping to syntactic structure, and suggests that it should be possible to find principled explanations for other such preferences, or tendencies, that have been observed.

References

Ackema, Peter, and Ad Neeleman. 2013. Person features and syncretism. *Natural Language & Linguistic Theory* 31(4): 901–950.

Ambrazas, Vytautas, ed. 1997. Lithuanian grammar. Vilnius: Baltos Lankos.

Bjorkman, Bronwyn, Elizabeth Cowper, Daniel Currie Hall, and Andrew Peters. 2019. Person and deixis in Heiltsuk pronouns. *Canadian Journal of Linguistics* 64(4): 574–591. doi: 10.1017/cnj.2019.13.

Bjorkman, Bronwyn M., Elizabeth Cowper, Daniel Currie Hall, Jennice Hinds, Louise Koren, and Dan Siddiqi. 2022. Morphological upstaging and 'markedness'. In *Proceedings of the 2021 annual conference of the Canadian Linguistic Association*, ed. Angélica Hernández and Chris

- Plyley, 13 pp. Canadian Linguistic Association.
- Bliss, Heather, and Elizabeth Ritter. 2009. A typological database of personal and demonstrative pronouns. In *The use of databases in cross-linguistic studies*, ed. Martin Everaert, Simon Musgrave, and Alexis Dimitriadis, 77–116. Berlin: Mouton de Gruyter.
- Bonet, Eulalia. 1991. Morphology after syntax: Pronominal clitics in Romance. Doctoral dissertation, Massachusetts Institute of Technology.
- Chomsky, Noam. 1995. The minimalist program. Cambridge, MA: MIT Press.
- Cowper, Elizabeth, and Daniel Currie Hall. 2014. Reductiō ad discrīmen: Where features come from. *Nordlyd* 41(2): 145–164.
- Cowper, Elizabeth, and Daniel Currie Hall. 2017. The rise of contrastive modality in English: A neoparametric account. *Linguistic Variation* 17(1): 68–97.
- Cowper, Elizabeth, and Daniel Currie Hall. 2019. Scope variation in contrastive hierarchies of morphosyntactic features. In *Variable properties in language: Their nature and acquisition*, ed. David W. Lightfoot and Jon Havenhill, 27–41. Washington: Georgetown University Press.
- Cowper, Elizabeth, and Daniel Currie Hall. 2022. Morphosemantic features in Universal Grammar: What we can learn from Marshallese pronouns and demonstratives. *Canadian Journal of Linguistics* 67(3): 242–266.
- Feldman, Harry. 1986. A grammar of Awtuw. Canberra: Australian National University.
- Greenberg, Joseph H. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton & Co.
- Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Ken Hale and Samuel J. Keyser, 111–176. Cambridge, MA: MIT Press.
- Halle, Morris, and Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21: 275–288.
- Harbour, Daniel. 2016. Impossible persons. Cambridge, MA: MIT Press.
- Harley, Heidi, and Rolf Noyer. 1999. Distributed morphology. Glot International 4(4): 3–9.
- Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3): 482–526.
- Ide, Sachiko. 1997. Joseego no sekai [The world of women's language]. Tokyo: Meiji Shoin.
- Janda, Laura A., and Charles E. Townsend. 2002. *Czech*. Durham, NC: Slavic and Eurasian Language Resource Center, Duke University. Published online at http://www.seelrc.org: 8080/grammar/pdf/compgrammar_czech.pdf.
- Kramer, Ruth. 2016. The location of gender features in the syntax. *Language and Linguistics Compass* 10(11): 661–677.
- Miyazaki, Ayumi. 2004. Japanese junior high school girls' and boys' first-person pronoun use and their social world. In *Japanese language, gender, and ideology: Cultural models and real people*, ed. Shigeko Okamoto and Janet S. Shibamoto Smith, 256–274. Oxford: Oxford University Press.
- Noyer, Robert Rolph. 1992. Features, positions and affixes in autonomous morphological structure. Doctoral dissertation, Massachusetts Institute of Technology.
- Plank, Frans, and Wolfgang Schellinger. 1997. The uneven distribution of genders over numbers: Greenberg nos. 37 and 45. *Linguistic Typology* 1(1): 53–101.
- Rath, John C. 1981. *A practical Heiltsuk-English dictionary with a grammatical introduction*. National Museum of Man Mercury Series. Ottawa: National Museums of Canada.

Ritter, Elizabeth. 1993. Where's gender? Linguistic Inquiry 24(4): 795-804.

Ritter, Elizabeth. 1995. On the syntactic category of pronouns and agreement. *Natural Language & Linguistic Theory* 13(3): 405–443.

Ritter, Elizabeth. 2014. Featuring animacy. Nordlyd 41(1): 103–124.

Siewierska, Anna. 2013. Gender distinctions in independent personal pronouns. In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL https://wals.info/chapter/44.