Morphological upstaging and "markedness"[*]

Bronwyn M. Bjorkman
Queen's University

Elizabeth Cowper
University of Toronto

Daniel Currie Hall
Saint Mary's University

Jennice Hinds
Carleton University

Louise Koren
Carleton University

Dan Siddiqi
Carleton University

Canadian Linguistic Association
June 4–7, 2021

1 Introduction

> Languages differ essentially in what they *must* convey and not in what they *may* convey.
>
> Jakobson (1959: 236)

Uncontroversially, languages make different choices about what grammatical information is morphologically obligatory.

- For example: both English and French mark finite verbs with both tense and subject agreement, while Mandarin exhibits no morphological inflection on verbs at all.

When a language **never** marks certain information morphologically—and is never syntactically sensitive to the distinction in question—we can safely analyze it as not representing the contrast in abstract representations.[1]

But in other cases, a language marks a distinction in some contexts, but that distinction is **systematically lost** in other contexts.

- For example: English systematically has no marking for **gender** on plural or participant pronouns.

Such neutralization is often asymmetric: one dimension of meaning is realized morphologically, while another is neutralized. We refer to this as **morphological upstaging**.

---

1. Though of course some linguists (e.g., Cinque and Rizzi 2008) have taken a *universalist* perspective, that distinctions made in any language are represented underlyingly in all languages.

In Distributed Morphology (DM; Halle and Marantz 1993 et seq.), upstaging involves competition between vocabulary items spelling out sets of features that are not obviously in a superset–subset relation.

For example, **person** and **number** each separately upstage **gender** in the Swedish pronoun system:

π	GENDER			SINGULAR	PLURAL
1	(HUMAN)			*jag*	*vi*
2	(HUMAN)			*du*	*ni*
3	HUMAN		MASCULINE	*han*	*de*
			FEMININE	*hon*	*de*
			UNSPECIFIED	*hen*	*de*
	NON-HUMAN		COMMON	*den*	*de*
			NEUTER	*det*	*de*

Table 1: Contemporary Swedish nominative pronouns[2]

- There are no gender contrasts in the first or second person, or in the plural.

- Suppose that the Swedish inventory of vocabulary items includes the following:

    - *du* ⇔ [PARTICIPANT]

    - *de* ⇔ [PLURAL]

    - *hon* ⇔ [FEMININE]

- When spelling out a pronoun with the features [PARTICIPANT, FEMININE] (to refer to a singular feminine addressee), why does *du* win over *hon*?
  ☞ This is second **person** upstaging **gender**.

- And when spelling out [FEMININE, PLURAL], why does *de* win over *hon*?
  ☞ This is plural **number** upstaging **gender**.

- We could stipulate that the 3SG VIs are marked for [−participant] and for [−plural], but that would seem to be missing a generalization, both within Swedish and typologically.

The existence of upstaging in the presence of potential "ties" of this kind raises two questions:

1. For a given pair of inflectional contrasts (e.g. NUMBER and GENDER), are patterns of upstaging cross-linguistically consistent?

2. Are these patterns, if they exist, illustrative of a key underlying principle in the synchronic grammar or are they better attributed to functional pressures or diachronic change?

---

2. For human referents, MASC. and FEM. reflect biosocial (so-called 'natural') gender. *Hen,* for a singular human referent of unspecified gender, is a relatively recent borrowing from Finnish *hän*; see Gustafsson Sendén et al. (2015). In the non-human 3SG pronouns, COMMON and NEUTER are grammatical genders. For 3PL, *de* is standard in writing, but *dom* is also common in speech. 2PL *ni* is also used as a polite 2SG.

Here we report on an ongoing study that investigates these questions in the nominal domain, specifically patterns of upstaging among **person**, **number**, **case**, and **gender** in pronominal and demonstrative systems.[3]

Our findings so far: There appear to be tendencies in which features are more likely to be upstaged, and which ones are more likely to upstage them, but no clearly categorical universals.

Outline:

2   Study

Languages included in this study were drawn from the database created as part of the survey in Bliss and Ritter (2009), as well as targeted studies of additional languages by two RAs on the project (Hinds and Koren).

2.1   Data from Bliss and Ritter (2009)

Bliss and Ritter (2009) created a database of pronominal and demonstrative systems in a typologically diverse range of languages.

- Every language included in the database was inspected for the presence of potentially-syncretic forms (forms that are surface identical).

- Since the database as it stood did not present paradigms in the usual chart format, and since we wanted to visually consider the paradigms for each language and the syncretisms they contained, we created a spreadsheet for each language in the database.

- The paradigms were organized in the familiar columns and rows, in the way that made the best sense for each language. We highlighted each syncretism in a distinct colour, as in Figure 1.

---

3. We use the term *gender* in a broad sense that includes noun classes in general, not just those with names like *feminine* and *masculine.*

Personal pronouns of Bandjalang

Form	Person	Case	Gender	Formality	General	Singular	Dual	Trial	Paucal	Plural
Stem	1st					ŋaj-				
Word	1st									ŋali
Stem	1st					ŋaj-				
Stem	2nd					wa-				
Word	2nd									ŋiri:maŋ
Word	2nd									ŋiri:maŋ
Stem	2nd					wa-				
Stem	2nd					wa-				
Word	3rd					njule				njulaŋam
Word	3rd					njule				njulaŋam
Word	3rd					njule				njulaŋam
Word	3rd					njule				njulaŋam

Bandjalang_exported.xlsx

A24 — Source: Geytenbeek, B., and H. Geytenbeek. 1971. Gidabal grammar and dictionary.

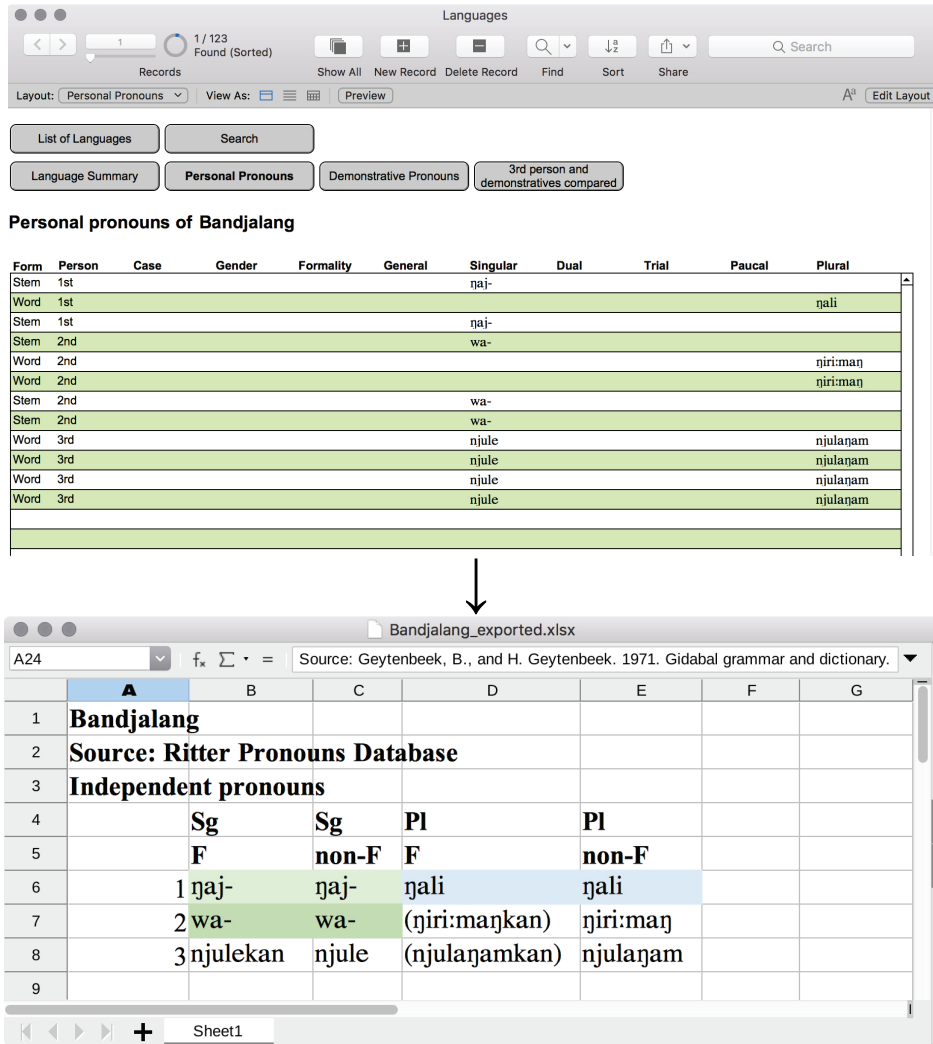|	A	B	C	D	E
1	**Bandjalang**				
2	**Source: Ritter Pronouns Database**				
3	**Independent pronouns**				
4		**Sg**	**Sg**	**Pl**	**Pl**
5		**F**	**non-F**	**F**	**non-F**
6	1	ŋaj-	ŋaj-	ŋali	ŋali
7	2	wa-	wa-	(ŋiri:maŋkan)	ŋiri:maŋ
8	3	njulekan	njule	(njulaŋamkan)	njulaŋam
9					

Figure 1: From database to spreadsheet: Bandjalang pronouns

2.2   Data from new survey

Independently, RAs on the project examine a smaller number of languages in detail.

- These languages do not represent a balanced typological sample. They were selected as languages likely to exhibit syncretism patterns that may involve upstaging, as well as based on availability of sources (given limited pandemic access to physical library resources).

- For each language, reference grammars were consulted to construct paradigms for pronouns, demonstratives, and (when relevant) agreement on adjectives and verbs.

- As with the paradigms from the Bliss and Ritter (2009) database, syncretisms were colourcoded for ease of later inspection.

- For each language a short report explaining the system of contrasts found in each paradigm was also written.

(1)   Languages with potentially relevant syncretisms identified so far
  a. Attic Greek   (Mastronarde 2013)
  b. Coast Tsimshian   (Dunn 1995; Forbes to appear)
  c. Hinuq   (Forker 2013)
  d. Levantine Arabic   (McLoughlin 1982)
  e. Modern Hebrew   (Glinert 1989)
  f. Slovene   (Herrity 2000)

2.3   Identifying upstaging

Once all instances of syncretism had been coded, we focused on syncretisms involving PERSON, NUMBER, CASE, and GENDER.

- Some languages had syncretisms involving politeness or distance (i.e. proximal vs. distal demonstratives).

- These were not examined further, and are not reported here, because there is less agreement in the literature about how these dimensions of meaning are encoded featurally.

- For languages where politeness involves a grammaticalization of number, the existence of a T/V politeness contrast (e.g., plural *vous* being used as a polite singular second person) was **not** counted as PERSON upstaging NUMBER.[4]

A syncretism was identified as involving **upstaging** only if a contrast present in one part of a paradigm was systematically absent in another column or row of the same paradigm.

- In some cases this determination was made on the basis of a review of the sources cited by Bliss and Ritter (2009), or other descriptions of the languages involved. The further sources consulted were: Demuth et al. (2009); Demuth (1988); Fennell and Gelsen (1980); Haas (1940); King (2003); Polomé (1967)

An important caveat: These findings are preliminary, based on others' descriptions of the languages, usually taken from grammars. A detailed morphosyntactic analysis of the features involved in each system could change our understanding of what's going on.

3   Results: Patterns of upstaging

The distribution of upstaging patterns in 78 languages examined so far is summarized in Table 2.

- Each cell gives the number of languages in which the relevant neutralization is found.

- Languages are counted in **each** cell for which they exhibit the relevant neutralization. The sum of all cells is thus more than 78.[5]

---

4. See Ritter and Wiltschko (2019) for an analysis of the T/V phenomenon based on 'recycling' rather than on the mechanisms by which other kinds of syncretisms are usually derived.

5. A version of this table with all language names appears in an appendix.

		category preserved			
		PERS	NUM	CASE	GEND
category upstaged	PERS		1	0	0
	NUM	12		3	7
	CASE	20	18		9
	GEND	28	38	14	

Table 2: Upstaging patterns in 78 language sample

In the remainder of this section we highlight a number of patterns within these overall results.

3.1 Which features are most likely to be preserved, and which neutralized?

The languages in our sample suggest a scale of features, with features higher on the scale more likely to be preserved at the expense of those lower on the scale in the spell-out of pronouns and demonstratives:

(2)   PERSON > NUMBER > CASE > GENDER

But these tendencies are not absolute:

- Person marking is much more likely to be preserved at the expense of other features. Additionally, person is almost never upstaged—but it is upstaged by number in the Navajo pronominal system. (Person also upstages the dual–plural contrast.)

π	SING.	DUAL	PLURAL
1	ʃí	nihí	tanihí
2	ní	nihí	tanihí
3	pí	taapí	taapí
4	hó	taahó	taahó

- **Number > person:** In the dual and plural, first- and second person forms are identical.

- **Person > number:** In the third and fourth persons, dual is not distinct from plural.

Table 3: Navajo personal pronouns[6]

- Gender is particularly likely to be upstaged by other features, and comparatively less likely to upstage anything itself—but it does upstage number 7 times and case 9 times in our sample.

- Number and case are somewhere in the middle.

---

6. The fourth person is used as something like an obviative (indicating a non-participant other than the one already designated by third person in the discourse), and also as a polite form for either a third person or an addressee (Reichard 1951: 81–82).

3.2 F > G and G > F

In some languages, we find both F upstaging G and G upstaging F **within a single paradigm**—*prima facie* evidence against universal or even language-specific rankings.

- In the Navajo pronouns in Table 3, non-singular number upstages the contrast between first and second person, but 3rd and 4th person upstage the dual–plural contrast.

- And in Swedish, Slovene, Polish, Lithuanian, Latin, German, and Attic Greek, case upstages gender **and** gender upstages case in the same paradigm.

CASE	MASC.	NEUT.	FEM.
NOM.	*òn*	*ôno*	*ôna*
ACC.	*njêga*	*njêga*	*njó*
GEN.	*njêga*	*njêga*	*njé*
LOC.	*njêm*	*njêm*	*njéj*
DAT.	*njêmu*	*njêmu*	*njéj*
INST.	*njím*	*njím*	*njó*

- **Case > gender:** The contrast between masculine and neuter is neutralized in all cases except the nominative.

- **Gender > case:** The contrasts between locative and dative, and between accusative and instrumental, are neutralized in the feminine.

Table 4: Slovene third-person singular pronouns (Herrity 2000)

- All of these languages are Indo-European—but this may be because Indo-European is over-represented among the languages with both gender and a rich case system.

3.3 Interactions of person and number with gender

There is a pattern of number and person conspiring to upstage gender.

- In 19 languages, both number and person upstage gender.

- In 16 of those languages, both upstagings are part of a single pattern.

  - In 11 of these languages, the gender contrast is only found in 3SG.
    Wolaytta (Table 5) is a typical example.

π	GEND.	SINGULAR			PLURAL		
		NOM.	ACC.	GEN.	NOM.	ACC.	GEN.
1		tani	tana	taga, taro	nuni	nuna	nuga
2		neni	nena	nega	inte	intena	intega
3 {	MASC.	i	a	ega	eti	eta	etana
	FEM.	a	o	iga	eti	eta	etana

- **Person > gender:** Gender is marked only in 3rd person.

- **Number > gender:** Gender is marked only in singular.

Table 5: Wolaytta personal pronouns (Ohman and Fulass 1976: 158)[7]

– Iraqw (Table 6) is unusual in having a gender contrast only in 2SG.

π	GENDER	SINGULAR	PLURAL
1		aníŋ	atén
2 MASC.	MASC.	kuúŋ	kuúngáʔ
2 FEM.	FEM.	kiíŋ	kuúngáʔ
3		inós	ino:ʔín

- **Person > gender:** Gender is marked only in 2nd person.

- **Number > gender:** Gender is marked only in singular.

Table 6: Iraqw independent subject pronouns (Nordbustad 1988: 30)

In 19 of the 28 languages where person upstages gender, this is because there is no gender contrast for participants—that is, the gender contrast is limited to third person.

3.4   A note on dual NUMBER

Of the 30 times NUM is upstaged, in 7 cases it is specifically **dual** number that is non-distinct from another value (usually plural, but occasionally singular).

- PERS upstages NUM only in loss of the dual in: Navajo, Tunica, Wappo
- CASE upstages NUM only in loss of the dual in: Slovene, Tonkawa
- GEND upstages NUM only in loss of the dual in: Ngandi, Tunica

4   Discussion: Explaining upstaging

Upstaging is a systematic pattern of syncretism in which a distinction present elsewhere in a language is neutralized in certain contexts.

- We can think of upstaging as a morphosyntactic counterpart of contextual neutralization in phonology.

- DM accounts for syncretism via either underspecification or from the application of language-specific morphological rules (i.e. Impoverishment).

- Upstaging is a puzzle for theories like DM because both the 'winning' realization and a competitor appear to be equally specified—something other factor seems to be needed to resolve the tie.

Is everything that looks like a tie really a tie?

- Many authors have linked syncretism to the co-occurrence of multiple 'marked' feature values (Jakobson 1939; Brøndal 1940; Greenberg 1966)

- But what it means for a feature value to be marked has remained in dispute—as has the question of whether the same features are marked in different languages.

---

7. Plural second- and third-person forms are also used as polite singulars. The choice between 1GEN.SG forms /taga/ and /taro/ is based on properties of the possessum.

- The **neo-parametric** view of formal features (Cowper and Hall 2017) allows for different languages specifying—syntactically 'marking'—different members of an opposition.

  - For two languages that contrast 1st and 2nd person, for example, it could be that one represents this as [SPEAKER] vs. $\varnothing$, while the other represents it as $\varnothing$ vs. [ADDRESSEE] (*pace* Harbour 2016).

What does this have to do with upstaging?

- Recall the vocabulary items proposed for Swedish in the introduction:

  - *du* $\Leftrightarrow$ [PARTICIPANT]
  - *de* $\Leftrightarrow$ [PLURAL]
  - *hon* $\Leftrightarrow$ [FEMININE]

- These vocabulary items—and the potential ties among them when spelling out feature combinations like [FEMININE,PLURAL]—assume that [PARTICIPANT], [PLURAL], and [FEMININE] are all specified features in Swedish.

- A single change—specifying [SINGULAR] rather than [PLURAL]–would substantially change the specification of the relevant vocabulary items, and the consequent competition among them.

For this reason, more careful morphosyntactic analysis of individual languages will be needed before we can conclude that the sample includes any true upstaging examples!

The context of typological universals:

Some of the tendencies we've identified in cross-linguistic patterns of upstaging are similar to generalizations proposed as implicational universals by Greenberg (1966).

- **Number > gender** is predicted by Universals 37 and 45:

  (3) **Universal 37.** A language never has more gender categories in non-singular numbers than in the singular (Greenberg 1966: 76).

  (4) **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).

- **Person > gender** (or at least *first* person over gender) is predicted by Universal 44:

  (5) **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).

But our results (so far) don't indicate a universal **hierarchy** of upstaging and upstageable features.

- Although we have no examples of case or gender upstaging person, for any other pair of features in the set we looked at, there's at least one apparent example of each upstaging the other.

5   Conclusion

Next steps…

- Continue survey of additional languages.

- Select languages from current survey for more detailed morphosyntactic analysis to determine relevant feature contrasts.

If this eliminates all cases of upstaging?

- An interesting result! (Though at this point we think it's unlikely.)

- Tells us something about possible and impossible morphological systems.

If we are left with some cases of upstaging?

- Investigating other sources for the preference for preserving some features over others.

- A tentative proposal: competition between vocabulary items may be constrained by a preference for realizing features with semantic content.

  – This preference predicts, e.g., the upstaging of (arbitrary) GENDER by NUMBER, and of (structural) CASE by ϕ-features.

  – But what this still does not account for is asymmetries between, say, PERSON and NUMBER.

- This is not (quite) a purely functional explanation: tied to a metafeature such as interpretability—thus (e.g.) the presence of lexical items expressing a given contrast will obviate its realization in the inflectional system.

Returning to questions posed in the introduction:

1. For a given pair of inflectional contrasts (e.g. NUMBER and GENDER), are patterns of upstaging cross-linguistically consistent?

   Yes. But these patterns are tendencies, not absolute.

2. Are these patterns, if they exist, illustrative of a key underlying principle in the synchronic grammar or are they better attributed to functional pressures or diachronic change?

   No answer yet. But clear directions for further investigation.

Thank you!

References

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.

Brøndal, Viggo. 1940. Compensation et variation, deux principes de linguistique generale. *Scientia* 9/10:101–109.

Cinque, Guglielmo, and Luigi Rizzi. 2008. The cartography of syntactic structures. *Studies in Linguistics* 2:42–58.

Cowper, Elizabeth, and Daniel Currie Hall. 2017. The rise of contrastive modality in English: A neoparametric account. *Linguistic Variation* 17:68–97.

Demuth, Katherine, 'Malillo Machobane, and Francina Moloi. 2009. Learning how to license null noun-class prefixes in Sesotho. *Language* 85:864–883.

Demuth, Katherine A. 1988. Noun classes and agreement in Sesotho acquisition. In *Agreement in natural languages: Approaches, theories and descriptions*, ed. M. Barlow and C. Ferguson, 305–321. Chicago: CSLI/University of Chicago Press.

Dunn, John Asher. 1995. *Sm'algyax: A reference dictionary and grammar for the Coast Tsimshian language*. Seattle: University of Washington Press.

Fennell, Trevor G., and Henry Gelsen. 1980. *A grammar of Modern Latvian*, volume 1. The Hague: Mouton.

Forbes, Clarissa. to appear. Tsimshianic. In *The languages and linguistics of Indigenous North America*, ed. Carmen Jany, Marianne Mithun, and Keren Rice, volume 1, 30 pp. Berlin: De Gruyter. Available on LingBuzz at `https://ling.auf.net/lingbuzz/004784`.

Forker, Diana. 2013. *A grammar of Hinuq*. Berlin: De Gruyter Mouton.

Glinert, Lewis. 1989. *The grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Greenberg, Joseph H. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton & Co.

Gustafsson Sendén, Marie, Emma A. Bäck, and Anna Lindqvist. 2015. Introducing a gender-neutral pronoun in a natural gender language: The influence of time on attitudes and behavior. *Frontiers in Psychology* 6:1–12.

Haas, Mary R. 1940. *Tunica*. New York: J.J. Augustin.

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.

Harbour, Daniel. 2016. *Impossible persons*. Number 74 in Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Herrity, Peter. 2000. *Slovene: A comprehensive grammar*. London and New York: Routledge.

Jakobson, Roman. 1939. Signe zéro. In *Mélanges de linguistique offerts à Charles Bally*, 143–152. Geneva: Georg et Cie, S.A.

Jakobson, Roman. 1959. On linguistic aspects of translation. In *On translation*, ed. Reuben A. Brower, 232–239. Cambridge, MA: Harvard University Press.

King, Gareth. 2003. *Modern Welsh: A comprehensive grammar*. London: Routledge, 2nd edition.

Mastronarde, Donald J. 2013. *Introduction to Attic Greek*. Berkeley, CA: University of California Press, 2nd edition.

McLoughlin, Leslie J. 1982. *Colloquial Arabic (Levantine)*. New York: Routledge.

Nordbustad, Frøydis. 1988. *Iraqw grammar: An analytics study of the Iraqw language*. Berlin: Dietrich Reimer.

Ohman, Walter A., and Hailu Fulass. 1976. Welamo. In *Language in Ethiopia*, ed. M. L. Bender, J. D. Bowen, R. L. Cooper, and C. A. Ferguson, 155–164. London: Oxford University Press.

Polomé, Edgar C. 1967. *Swahili language handbook*. Washington: Center for Applied Linguistics.

Reichard, Gladys A. 1951. *Navaho grammar*. New York: J.J. Augustin.

Ritter, Elizabeth, and Martina Wiltschko. 2019. The syntax of formality: Universals and variation. Presented at the 2019 annual meeting of the Canadian Linguistic Association, University of British Columbia, June 2019.

A   Language names in results

| | CATEGORY PRESERVED | | | |
	PERS	NUM	CASE	GEND
PERS		**1:** *Navajo*	**0**	**0**
NUM	**12:** *Berik, Coast Tsimshian, Haitian Creole, Koasati, Kutenai, Kwakiutl, Maxakalí, Navajo, Tauya, Tunica, Wappo, Wichita*		**3:** *Balochi, Slovene, Tonkawa*	**7:** *Arapesh, Catalan, Ho, Ngandi, Pakaàsnovos, Sotho, Tunica*
CASE	**20:** *Albanian, Catalan, Coast Tsimshian, Dutch, Fijian, Georgian, German, Godie, Greek, Hausa, Hinuq, Kabardian, Kwakiutl, Latin, Pidgin Nigerian, Polish, Spanish, Tamazight, Wappo, Yupik*	**18:** *Albanian, Attic Greek, Catalan, Coast Tsimshian, Comanche, Dutch, Georgian, German, Hausa, Latin, Lithuanian, Pidgin Nigerian, Slovene, Spanish, Wappo, Xokleng, Yaoure, Zuni*		**9:** *Attic Greek, German, Hausa, Latin, Lithuanian, Polish, Swedish, Telugu, Xokleng*
GEND	**28:** *Arabic (Gulf), Arabic (Levantine), Awtuw, Bandjalang, Catalan, Cuebo, Dieri, Ho, Iraqw, Latin, Latvian, Modern Hebrew, Nama, Ngandi, Pakaànovos, Pomo, Rikbaktsa, Salish (Southern Puget Sound), Somali, Sotho, Spanish, Swedish, Tamazight, Tunica, Welsh, Wolaytta, Xokleng, Yimas*	**38:** *Albanian, Arabic (Gulf), Arabic (Levantine), Arapesh, Attic Greek, Awtuw, Cubeo, Dieri, Dutch, German, Godie, Halkomelem, Hausa, Hinuq, Iraqw, Kannada, Latin, Lithuanian, Luiseño, Miwok (Central Sierra), Modern Hebrew, Nama, Ngandi, Pakaàsnovos, Polish, Pomo, Rikbaktsa, Romanian, Salish (Southern Puget Sound), Slovene, Somali, Sotho, Swahili, Swedish, Telugu, Welsh, Wolaytta, Xokleng, Yimas*	**14:** *Albanian, Attic Greek, Djingili, Georgian, Hinuq, Latin, Latvian, Lithuanian, Luiseño, Polish, Romanian, Slovene, Swedish, Tamazight*	

(rows labelled vertically: CATEGORY UPSTAGED)

Table 7: Upstaging patterns in 78-language sample