**Competing Motivations for Gender Distinctions
Relative to Numbers in Pronominal Paradigms:
Markedness, Conflict Resolution, Economy, Inertia
— and Ways of Knowing: Which Will Win?**

Frans Plank
(Universität Konstanz)

(1) Universal 37
   A <u>language</u> **never** has more gender categories in nonsingular numbers (dual, ambal, trial, quattral, paucal, multal, plural, augmented, associative, ...) than in the singular (unit).

(2) Universal 45
   If there are any gender distinctions in the plural (or other non-singular numbers) of the <u>pronoun</u>, there **always** are some gender distinctions in the singular (unit) also.
   (after Greenberg 1963, inter multos alios)

(3) Universals exegesis

   What about Spanish personal pronouns?
   Or genitive of pronominal inflection in Latin?
   Or (possible?) languages where singular genders are Masculine, Feminine, Vegetable, Other and plural genders are Animate and Inanimate?
   Or languages where singular is uniformly zero and plural has an overt exponent only for nouns/pronouns high on the animacy scale (and there is no agreement in gender/class)?

   • What **are** genders/classes and how to distinguish them?
   • Are the **same** genders required for SG and non-SG?
   • Is the universal supposed to hold for **languages**, **words**, or **word forms**?

(4) Universal 37/45 in strictest form
   If there are any gender distinctions in any non-singular numbers of any inflected word form of any language, then there will always be the same or more gender distinctions in the corresponding singular form of this word form also.

(5) The Facts
   Non-singulars richer in gender than singulars, even such as are only proscribed by the more liberal interpretations of the relevant universal, do exist — with a disturbingly high incidence of up to 25% in the languages examined (n=400+)
   (Plank & Schellinger 1997).

   for example:

   (i)   transparently adding to old pronouns:
            Spanish (IE),
            Lithuanian, Slovene (IE), Hayu (Tib-Burman),

Kamoro (Trans-New Guinea),
Vanimo (Sko)

(ii)   unaccountably deviant duals and trials (or also unit augmented), with more
genders than PL, sometimes also than SG:
Murui Witoto (Macro-Carib),
Baniata, Lavukaleve (Yele-Solomons, East Papuan),
Au, Olo (Torricelli),
some non-Pama-Nyungan (Australian)

(iii)   3rd PL/non-SG only: specifically animate/human PL forms:
Biak, Windesi/Wandamen (West NG, Austronesian)
Katu (Mon-Khmer), Palau (Austronesian)

(iv)   not 1st SG:
Ntifa et al. (Berber), Hadza, Nama (Khoisan)

(v)   (almost) anything goes — in Berber

(6) **What does this mean?**

(7) **Congratulations!**

"Statistical universals are better than absolute universals" (Dryer 1997):
They are not only "falsifiable" but "testable", i.e., you can show that they are **true**,
which you can't for an absolute universal, which would have to be true of **all possible**
human languages.

However, there remains the problem of how to tell whether unequal crosslinguistic
frequencies, though in line with a statistical universal, really reflect **structural
preferences** or ar due to **historical accidents** (population movements, language shifts)
or are even **random variations**. But then, these are also possibilities for absolute
universals (massive ignorance, mass non-documentation, mass decimation).

(8) **Too bad! But can't we do better?**
Yes: Ersatz Universal 45, in four parts
(p authorizes q = no q without p)

(i)   A gender distinction in the singular authorizes the same gender distinction in
non-singular numbers of the corresponding person. (≈ Universals 37/45 strict)
(ii)   A gender distinction in 3rd person singular authorizes the same gender
distinction in 1st and/or 2nd person non-singular.
(iii)   A gender distinction in 3rd person non-singular needs no special authorization
from gender elsewhere.
(iv)   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.

(9) Cf. this one (Dryer 1997):
If basic word order is VO, then **always** N RelCl.
Counterexample to this absolute universal: the Chinese languages.
Accordingly revised, now without counterexamples:
If basic word order is VO and there are no tones, then **always** N RelCl.

But: Although probably a valid descriptive generalization, is this "a significant observation about language"?

(10) **Scrap universal, ersatzlos**

Find **motivations** for the various attested distributions of gender distinctions relative to numbers, — and maybe there are equally good structural or functional reasons for distributing them one way or another

(11) **Markedness** (cognitive complexity, formal economy)

When for category A not all terms are equally well distinguished in the paradigmatic company of all terms of (dominant) category B, then it will be in the company of **marked** term(s) of category B that fewer terms of category A will be distinguished.

Thus, if asymmetries:
fewer genders in (marked) non-singulars than in (unmarked) singulars, and i(marked) oblique than in (unmarked) direct cases;
fewer cases in non-singulars than in singulars;
fewer persons/numbers in (marked) past than in (unmarked) present, and in (marked) subjunctive than in (unmarked) indicative;
etc., etc.

(12) Strategies of avoiding or **resolving gender conflicts** in non-singular anaphora or agreeement
    he and he ... he.PL;
    she and she ... she.PL;
    he and she ... ?.PL

   (i)   designating one of the existing genders (the unmarked one) in non-singular as the resolution gender
            (e.g. French *il et elle ... ils*);

   (ii)   distinguishing no genders in non-singular
            (e.g. English *he and she ... they*);

   (iii)   providing an **extra** mixed gender in non-singular
            (e.g., Nama, !Xû, widespread also elsewhere in Khoisan, universal in Central Khoisan; Bukiyip [Torricelli], Vanimo [Sko], and perhaps other "Papuan"; probably Turkana [Nilo-Saharan], Khasi [Mon-Khmer], and Chinook [Chinookan]).

   (i) and (ii) are more **economical** than (iii) in terms of the number of forms required to resolve gender conflicts;
   (i) might be considered more complex than (ii) and (iii) since a special **resolution rule** is required, but it can probably be motivated more generally in terms of markedness.

(13) **Inertia**: Retention of (especially recent) inheritance

Once forms which happen to distinguish genders (such as numerals, quantifiers, adjectives, nouns) have been grammaticalized in pronominal paradigms to add or accentuate non-singular numbers (and it is typically non-singulars rather than singulars which are thus innovated or renovated!), why change more and, at the earliest convenience, discontinue such recently or not so-recently imported formal distinctions?

for example:
-otros/-otras 'others' in Spanish,
-du/-dvi 'two' in Lithuanian,
-dva/-dve 'two' in Slovene,
-hei/-bei 'he'/'she' in Vanimo,
kaᵒka 'woman' in Kamoro

(14) Given three possible motivations, why are patterns conforming to Markedness/Economical-Conflict-Resolution **so much more frequent** than those conforming to Conflict-Resolution-By-Extra-Form or Inertia-After-Innovation?

(15) Are they?
— probably, on the evidence of **The (Relatively) Isolates Sample** (smallish, but maximal genetic independence guaranteed):

(i) *symmetrical: no gender in any number*	(ii) *asymmetrical: fewer genders in non-SG*	(iii) *asymmetrical: more genders in non-SG*	(iv) *symmetrical: same genders in all numbers*
Nivkh	(Basque?)	---	Basque
Hunza Burushaski	Ket		Hurrian
Kutenai	Sumerian		Yasin Burushaski
Zuñi	Hattic		Naháli
	(Hunza Burushaski)		Ainu
	Yuchi		Tunica
	Tiwi		

(16) Chance (random variation, historical accidents — Indo-Europeans as the colonial powers rather than, say, the Khoisan) or (structural) necessity?

(17) For lots of other rather reasonable things, or also some rather unreasonable ones, which are rare or indeed nonesuch, see *Das grammatische Raritätenkabinett* at
http://ling.uni-konstanz.de/pages/proj/sprachbau/htm

(18) The relationship between typology and diachrony:
Who is in charge?

(i)   Universals impose limits on variation across languages at any and all times; they thereby constrain change insofar as languages cannot change so as to violate a timeless (genetic or functional) law, or at any rate not without subsequent changes redressing the balance one way or another.
(Possibly: There are no laws of change itself.)

(ii)   Assuming that particular targets (forms, categories, constructions, rules, constraints, etc.) can only result from particular mechanisms of change (reanalysis) operating on particular source forms (categories, etc.), this would impose limits on how languages can differ: they can only be what they could become.
(Possibly: There are no timeless laws.
Or: Concomitant or consecutive changes are superintended by timeless laws.)

(19) infixes ⊃ adfixes,
      covering law:
      discontinuous constituents ⊃ continuous constituents,
      which, so much for a cognitive explanation of the descriptive generalization, are easier to
      store and process

(20) infixes < adfixes,
      the only source, and only two mechanisms of change: phonological **metathesis** to
      optimize syllable structure (and possibly subsequent morphologization) or **entrapment**
      (outer adfix reanalysed as part of stem)

(21) overt indefinite forms (articles, pronouns) in plural NPs
                        ⊃ overt indefinite forms in singular NPs,
      explanation: markedness reversal, individuation being marked in "ignorative" contexts
      (indefinite, interrogative) and therefore requiring some extra formal expenditure

(22) indefinites   <   numeral 'one' (dedicated singular),
                                    mid-range quantifiers (dedicated plural),
                                    interrogative pronoun,
                                    generic noun,
                                    ...

                        by grammaticalization,
                        and possibly subsequent extensions of number distinctions

(23) diachronically, overgeneration is possible, if change is left to itself:
      *SG <u>u</u> indef   PL <u>m</u> indef;
      prevented by timeless constraint, superintending concomitant reanalyses:
      Grammaticalize and analogically extend as is permitted (or indeed required) by
      diachronic law,
      but don't create an overt plural indefinite form unless there already is a singular
      indefinite form.

(24) The universals about gender distributions over numbers at (1), (2), (4), and (8) above are
      conceived of as **timeless laws**, superintending grammar construction (paradigm
      structuring) at any and all times.

      Can they be replaced or complemented by **diachronic laws**, thereby avoiding the
      conclusion that the patterns actually found, some more frequent than others, are random
      structural variations or due to historical (non-structural) accidents?

(25) Sources of gender distinctions in (personal and demonstrative) pronouns, and mechanisms of
      their innovation, alteration (extension, reduction), and loss:

      (i)   gender-distinguishing words (classifiers, generic or relational nouns, distal-deictic
               demonstratives or adverbs, copulas) used or reanalysed as personal pronouns
               (e.g., *it – there* contrast in 3rd person in English);

      (ii)   gender-distinguishing words (numerals, quantifiers, adjectives, nouns) attached to
               existing personal pronouns
               (e.g., hypothetically, *you-guys* 2PL.MASC – *you-girls* 2PL.FEM);

(iii)   gender-distinguishing singular pronoun used as a constituent of plural pronouns
(e.g., *we-he* 1EXCL.MASC – *we-she* 1EXCL.FEM; *he-and-she* 3PL)
[cf. (8ii) of the authorization version of the Ersatz universal above];

(iv)   gender distinction analogically extended from other pronouns of the same number or across numbers (or also from nouns);

(v)   borrowing of gender-distinguishing pronouns;

(vi)   gender distinction obliterated phonologically;

(vi)   gender distinction levelled out morphologically;

(vii)   gender-distinctive pronoun(s) discontinued.

(26)   While it is in fact typical for relevant innovations and renovations (especially per (ii)) to be limited to non-singular parts of pronominal paradigms (with singulars being diachronically rather stable), thereby selectively introducing gender distinctions when the source forms themselves distinguish gender and such formal contrasts are transparently retained, sources and mechanisms for creating **mixed-gender** forms are not so obvious. Possible candidates are indefinites reanalysed as (definitish) personal pronouns (as has been argued for Khoisan) and default plural gender forms for inanimates extended to animates (as in Bukiyip). Given that such mixed-gender forms are unlikely products of change, and hence are also unlikely to be available for crosslinguistic propagation through historical accidents of relevant languages becoming numerically predominant, why legislate against them through a timeless law superintending change?

(27)   (i)   DEF reanalysed/used as INDEF
3PL   'they have burgled our house again'
2SG/PL   'you never know'
3SG   unmarked pro filling argument positions
(ii)   INDEF reanalysed/used as DEF
only mixed genders?

(28)   The rarer, the transienter?

(i)   How long does it take to **introduce** (= to innovate and get accepted) X?
(ii)   What is the **life expectancy** of X
— when left to itself, when in competition with Y?
(iii)   How long does it take to **discontinue** X?

X, Y = anything linguistic, e.g.
a feature, a phoneme,
a paradigmatic distinction or pattern, a category,
a word, a morphological exponent, a construction,
a rule, a constraint, an ensemble of rules/constraints,
a language/grammar&lexicon;

linguistic time measure:
number of generations of grammar acquirers/re-formers
(one generation = ca. 30 years, or less in the old days;
thus: 200,000 years = ca. 8.000 generations, or cycles of language-acquisition)

Abstract

Asymmetric skewings of paradigmatic distinctions in subparts of pronominal paradigms are widely believed to follow universal patterns, which are generally assumed to be explicable in terms of markedness, itself explicable in terms of formal and cognitive complexity. Thus, one well known descriptive generalization is this: There are never more gender (including class) distinctions in non-singular numbers than in the singular of personal pronouns. The standard explanation then is that this is because quattral, trial, dual, and plural are marked numbers vis-a-vis singular, and the pressure to economize on forms, resulting in inflectional homonymy of gender, will be felt the more strongly the more cognitively complex the categorial partnership of gender.

Actually, as has been shown recently (Plank & Schellinger in LT 1, 1997), this purportedly absolute universal about gender distinctions relative to numbers, in pronominal paradigms and elsewhere, is not really universally valid, but to only about 75%: about one out of four languages (or somewhat fewer, depending on one's notion of gender and one's strictness in requiring gender distinctions in corresponding singular and non-singular persons) is not well-behaved, offending against markedness considerations. There is diversity rather than uniformity on this count, then.

Moreover, there are at least two very good reasons why pronominal paradigms may offend against markedness principles. First, in order to resolve agreement conflicts, it makes as much sense to have an extra mixed gender in non-singulars (as in Nama, Vanimo, Bukiyip; schematically: *he and he … he-s, she and she … she-s, he and she … chee-s*) as to have either no gender distinction in non-singulars (as in English: *he and she … they*) or some gender-ranking strategy (as in French: *il et elle … ils*). Second, non-singular forms may only recently have been grammaticalized from words richly inflecting for gender (such as adjectives, numerals, quantifiers, as e.g. in Spanish plurals *nos-otros/nos-otras* 'we (MASC/FEM)' or Lithuanian duals *mù-du/mù-dvi* 'we two (MASC/FEM)')—and why should these gender distinctions be abolished by the next generations at the earliest convenience?

This raises the question why of three alternative patterns of paradigmatic skewing, each well motivated, one prevails far more frequently than the two others, albeit not to their universal exclusion. One possible answer is that the prevalence of markedness over conflict resolution and the retention of recent inheritance in determining asymmetric distributions of genders over numbers is due to chance rather than necessity, as far as grammatical structure itself is concerned. Structural diversity, to the extent that it is regular, is shaped (i) by the timeless implicational laws of typology and (ii) the historical laws governing the restructuring of grammars by successive generations of learners, including those learning more languages than one. In the case at issue, diversity does not seem a matter of necessity in either sense. If all grammatical laws fail, timeless (typological) or timeful (diachronic), what remains to explain unequal crosslinguistic distributions of non-universal grammatical traits is chance—the good or bad luck of being widely or not so widely spread as peoples move around, get in touch, cause others to discontinue their own grammars, or perhaps get wiped out themselves.

One way to hear this paper, then, is as a plea for typology to be done in concert with history—that of grammars and of peoples.

Frans Plank
Sprachwissenschaft
Universität Konstanz
78457 Konstanz
Germany
frans.plank@uni-konstanz.de