136.–137. Personal Pronouns

Johanna Nichols and David A. Peterson

1.   Introduction

Personal pronouns, like other closed-class forms, tend to contain nasals and other basic consonants. However, particular paradigmatic combinations of consonants in particular person-number categories are not especially frequent worldwide. There are only two such combinations that show up interestingly on a map: pronoun sets like English *me* and *thee* or Nanai (Tungusic; eastern Siberia) *mi* 'I, me' and *si* 'you', with *m* in the first person and a coronal obstruent in the second, which are common in northern Eurasia and rare elsewhere; and sets like Northeast Maidu (Penutian; California) *ni* 'I, me' and *mi* 'you' with *n* in the first person and *m* in the second, which are common in western America and rare elsewhere. Map 136 here shows the first type, henceforth *"m-T"* patterns, and Map 137 the second, henceforth *"n-m"* patterns.

The maps here are based on the first and second person singular forms only. They survey the first consonant in the pronoun root, whether it is word-initial as in Wintu (Penutian; California) *ni* 'I, me; we, us' or follows an initial vowel as in Nez Perce (Sahaptian; Idaho and Oregon) *'iin* 'I, me'. (Whether written or not, the glottal stop that precedes an initial vowel is not considered the first consonant. The Nez Perce form just cited begins with such a glottal stop.) They survey three pronominal forms: **independent pronoun** words like English *I, me, you*; **possessive affixes** as in Hungarian *ház-am* 'my house', *ház-ad* 'your house'; and **verbal affixes** as in Rama (Chibchan; Nicaragua) *n-taak-i* 'I am going', *m-taak-i* 'you are going' (Grinevald n.d.). (Of course, not all languages have all three series.) They consider the two most widely distributed

allomorphs of each, e.g. English *I* and *me* for first person singular.

2. Defining "m", "n" and "T"

The consonants mapped are defined as follows:

> *m* = [m] as well as glottalized, voiceless, lenis, and fortis labial nasals (in fact only plain [m] occurred in the language sample)
>
> *n* = dental or alveolar [n], palatal [n´] (and glottalized, voiceless, etc. variants, though none appeared in the sample)
>
> *T* = any apical obstruent: [t], [d], [s]; palatals such as [č], [š], etc.

Here are some examples of pronouns that fit the two types. First-person *m* and second-person apical obstruent in independent forms:

(1)   German
        1SG   *mich* (accusative)
                mein (possessive)
        2SG   *dich* (accusative)
                dein (possessive)

(2)   Georgian
        1SG   *me*
        2SG   *šen*

(3)   Nanai (Tungusic; eastern Siberia)
        1SG   *mi*
        2SG   *si*

in possessive affixes:

(4)   Hungarian
     1SG   *ház-am*   'my house'
     2SG   *ház-ad*   'your house'

in verbal affixes:

(5)   Turkish present tense (Lewis 1967: 109)
     1SG   *alıyor-um*   'I am taking'
     2SG   *alıyor-sun*   'you are taking'

First person *n* and second person *m* in independent forms:

(6)   Wintu (Penutian; California)
     1st person   *ni*
     2nd person   *mi*

in possessive affixes:

(7)   Pipil (Uto-Aztecan; El Salvador; Campbell 1985: 43)
     1SG   *nu-chi:l*   'my chili pepper'
     2SG   *mu-chi:l*   'your chili pepper'

(8)   Karok (Hokan; California; Bright 1957: 560)
     1SG   *nani-tta:t*   'my mother'
     2SG   *mi-tta:t*   'your mother'

in verbal agreement formatives:

(9)   Nisgha (Tsimshianic; British Columbia; Tarpent 1987: 461)
     1SG   *nə-*
     2SG   *mə-*
     (transitive subject proclitics)

3. Defining the map values

The following values are shown on the maps:

Map 136: M–T Pronouns

@	1.	No M–T pronouns		200
@	2.	M–T pronouns, paradigmatic		27
@	3.	M–T pronouns, non-paradigmatic		3
			total	230

Map 137: N–M Pronouns

@	1.	No N–M pronouns		194
@	2.	N–M pronouns, paradigmatic		25
@	3.	N–M pronouns, non-paradigmatic		11
			total	230

The maps distinguish between **paradigmatic** and **non-paradigmatic** distributions of the consonants. The consonants form a paradigm when they both occur in the same form class(es) of their respective pronouns. For instance, a consistent paradigm occurs in Mapudungun (Araucanian; Chile):

(10)   Mapudungun (Smeets 1989)

		1SG	2SG
Independent forms:	Nominative	*iñche*	*eymi*
	Possessive	*ñi*	*mi*
Verbal suffixes:	Indicative	*-(ü)n*	*-m-i*
	Conditional	*-i-Ø*	*-m-i*

The underlined consonants are in the same position in three different form series and occur in both forms of the set. A subparadigm occurs in Wichí (Mataco-Guaicuru; Argentina):

(11)   Wichí (Viñas Urquiza 1974)

		1SG	2SG
Independent Nominative		*oλam*	*am*
Possessive	animate	*olä*	*alä*
	inanimate	*okä*	*akä*
Verb subject		*o–*	*le–*
Verb object		*–no*	*–am*
Verb object		*–o*	*–a*

The first pair of verb–object suffixes (the next to last line here) shows the pattern; no other does. The maps in this chapter do not distinguish consistent paradigms from subparadigms.

A case of first person *n* and second person *m* which do not form a paradigm occurs in Asmat (Asmat–Kamoro; Papua, Indonesia):

(12) Asmat (Voorhoeve 1965b)

	1SG	2SG
Independent Nominative	*no–r*	*o–r*
Independent Oblique	*no*	*o*
Verb subject	*–i*	*–em*
Verb object	*–(e)n*	*–(e)n*

Here there is first person *n* in the independent forms and second person *m* in the verb–subject agreement suffix. In fact, *n* occurs in both persons in the verb object forms. There is no place where *n* and *m* cooccur in the same form class.

Finally, the occasional token of one or the other paradigm can turn out to be truly artifactual, as in the following from Grebo (Kru, Niger–Congo; Liberia):

(13) Grebo (Innes 1966)

	1SG	2SG
Emphatic	*mɔ 1*	*mɔ 2/3*
Object	*mo 1/2*	*mo 2/3*
Possessive	*na 1/2*	*na 3*

(numbers indicate tones)

The consonant-vowel forms of the pronouns are the same in all form classes, and only the tones differentiate them. One can find a first person *n* and second person *m* in this set, so this technically counts as a (non-paradigmatic) occurrence of the *n-m* consonantism, but this is rather clearly due to a fluke rather than to a principled linguistic analysis of the material.

It is the paradigmatic sets that are most interesting, and in both maps they are by far the most common of the examples of the two consonant pairs. Non-paradigms (like that of Asmat, above) and flukes (like the forms of Grebo just above) are rare, and occur only outside of the areas where the paradigms are relatively common.

4.   Geographical distribution

The pronoun systems in question form two large areal clusters: an *m-T* one centered in northern Eurasia (Map 136) and an *n-m* one extending from western coastal North America to western South America (Map 137). There are a number of other maps in this atlas that have similar areal distributions: the *n-m* concentration on the Pacific Rim and especially the American Pacific Rim recalls the distribution of multiple possessive classification, bound inalienables, numeral classifiers, verb-initial word order, and double marking (see chapters 59, 58, 55, 81, 23, respectively); the *m-T* cluster in northern Eurasia recalls other Greater Silk Road distributions such as dependent marking, low synthesis, absence of possessive classification, and case-number coexponence (see chapters 23, 22, 59, 21, respectively).

5.   The inset maps

Mapping just *m* in either first person singular (Map 136A) or second person singular (Map 137A) yields somewhat more diffuse clusters but reveals the same large areas clearly.

@	1.	No *m* in 1st person singular		177
@	2.	*M* in 1st person singular		53
			total	230

Values of Map 136A. M in 1st Person Singular

[Map 136A about here]

@	1.	No *m* in 2nd person singular		152
@	2.	*M* in 2nd person singular		78
			total	230

Values of Map 137A. M in 2nd Person Singular

[Map 137A about here]

First person *m* (Map 136A) is ubiquitous in the Greater Silk Road area and nearly pan-Eurasian, missing in the Eurasian islands, Southeast Asia, and the Asian Pacific Rim; it is well attested in sub-Saharan Africa and New Guinea, sparsely attested in the Americas, and absolutely absent from Australia. Second person *m* (Map 137A) has a solid Pacific Rim distribution with occasional tokens inland in the Americas, in the Eurasian islands, and in Africa; it is absolutely absent from the Greater Silk Road area and Australia.

6. Explanations

What explains these two large unique clusters of languages with the distinctive pronoun consonantisms? The two main explanations that have been offered in the literature are:

· Sound-symbolic universals: nasals and other basic consonants are the most elementary consonants, hence the

most easily learned by children, and such consonants figure frequently in pronouns and other basic words. While this generalization is correct for broad phonological categories like "nasal", it cannot explain the distributions of the two-consonant paradigms or even of just *m* mapped here, for their distribution is not universal. No explanation based on universals can account for the fact that each of the two paradigms is fairly common in one large area and vanishingly rare elsewhere. The near-absence of each paradigm in every large area but one indicates that they are not due to any underlying tendency but have particular historical origins.

Genealogical relatedness of the languages: each cluster is made up of the descendants of a single protolanguage. While this explanation may seem plausible, there is no other evidence for the relatedness of all and only the conforming languages, and the pronoun systems are not sufficient evidence to prove relatedness. The conforming languages in each area belong to several different well-established families:

m–T:

Language(s)	Family
Eurasia:	
Chukchi, Itelmen	Chukotko-Kamchatkan
Chuvash, Turkish, Tuvan	Turkic
Finnish, Hungarian, Nenets	Uralic
French, German, Modern Greek, Hindi, Ossetic, Persian, Russian, Spanish, Waigali	Indo-European
Georgian	Kartvelian
Khalkha	Mongolic
Nanai	Tungusic
Kolyma Yukaghir	(isolate)
Elsewhere:	
Nigerian Fula, Grebo	Niger-Congo
Lakhota	Siouan

Salt–Yui	Chimbu
Southern Sierra Miwok	Utian
Usan	Adelbert Range

n–m:

Language(s)	Family
Americas:	
Axininca Campa	Arawakan
Cashinahua	Panoan
Chimariko	(isolate)
Upper Chinook	Chinookan
Highland Chontal	Tequistlatecan
Mesa Grande Diegueño	Yuman
Kiowa	Kiowa–Tanoan
Luiseño, O'odham, Pipil, Yaqui	Uto–Aztecan
Northeast Maidu	Maiduan
Mapudungun	Araucanian
Miskito	Misumalpan
Mixe	Mixe–Zoque
Rama	Chibchan
Northern Sahaptin	Sahaptian
Wari'	Chapacura–Wanhan
Wichí	Matacoan
Wintu	Wintuan
Yawelmani	Penutian
Elsewhere:	
Dumo	Sko
Supyire	Niger–Congo

On the other hand, in each cluster there are a few language families that are likely to be deeply related. Indo–European and Uralic are likely to be related, though the relatedness is so distant that we may never have sufficiently firm evidence to prove it by conventional means (e.g. Ringe 1998, Oswalt 1991). Still deeper relatedness of several of the other *m–T* families of

Eurasia to these has been proposed but never demonstrated. In any event the first person *m* is secondary in Turkic and Tungusic (where it developed from a *b*). In North America, a set of language families which have the *n–m* system is thought to be possibly deeply related: Utian, Wintun, Maidun, Sahaptian, and possibly also Chinookan may belong to a proposed megafamily called Penutian. But not all of the putative Penutian daughter families exhibit the system, and the putative Penutian set represents less than half of the tokens of *n–m* systems. Furthermore, one putative Penutian language, Southern Sierra Miwok, exhibits the *m–T* system. Thus deep genetic relatedness is no explanation for the pronoun patterns.

Both the *m–T* and the *n–m* paradigms are found in areas where spreads are known to have been centered -- the Greater Silk Road and the Pacific Rim – and both appear to be the products of geographical spread rather than just universals or just inheritance. Such spreads of pronominal consonantism patterns evidently do not occur often: we have only these two clear cases in the whole world. The origins of both are old (older than any of the individual families exhibiting the systems, for instance): the Eurasian *m–T* pattern is at least pre-Neolithic (Nichols 2001), and the American *n–m* one is late glacial at the latest (Nichols and Peterson 1996). Since these ages are greater than the time depth to which linguistics can generally trace genealogical relations, the details of their origins are lost in time.

We also looked for other interesting consonant patterns in pronominal roots, e.g. those involving velar nasals and velar obstruents *k, g,* etc. However, no such patterns showed conspicuous large areal distributions; all are more or less evenly distributed worldwide. So far, then, the *m–T* and *n–m* paradigms mapped here are unto themselves.

For more discussion see Nichols and Peterson 1996 and references therein; for subsequent discussion, see Nichols 2001.