

Taxonomy, Description, Definition, Explanation: Special Case:
Pronouns

Author(s): David Justice

Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society (1978), pp. 96-105

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Taxonomy, Description, Definition, Explanation:
Special Case: Pronouns [excerpt]

David Justice
U.C. Berkeley

The question of general interest behind the specific discussion that follows, which unfortunately cannot be here answered in a general way, is what is the point of labels and categories in the study of language. [...] At the simplest level, getting clear on the status of our labels will permit us to avoid writing such passages as the following (from a standard history of French): "The important role played by prepositions in Modern French as compared with Classical Latin is largely the result of the general analytical trend of the language." 'The general analytic trend of the language' is in fact not defined independently of, precisely, such phenomena as the replacing of Latin synthetic forms by prepositional phrases.

The analysis of pronoun systems has been attractive not only to linguists but to anthropologists, whom one would expect to have particular concern for avoiding sterile hypostasizing and for cashing their constructs in reality. Yet the feature-analyses offered may be given with only the most abstract motivation.

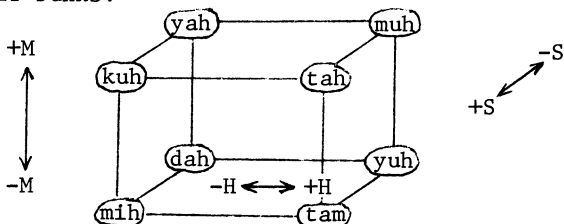
One criterion that some analysts (of phonemic and kinship systems as well as of pronoun systems) have favored for a satisfying analysis is high, preferably exhaustive occupancy of the cubby-holes established by intersection of the features one chooses. That is, given features $\pm F_1, \dots, \pm F_n$, one would ideally like to see all combinations realized by some linguistic form. If not, the system is 'asymmetrical', or has gaps.

Harold Conklin, in a well-known article ("Lexicographic treatment of folk taxonomies", *IJAL*, 1962), considers Hanunóo pronouns dah "they"; kuh "I"; mih "we (exclusive)"; muh "you (sg.)"; tah "we two (inclusive)"; tam "we-all (inclusive)"; yah "he, she"; yuh "you-all" and rejects the traditional labels "1, 2, 3 person". That is, he rejects them as formal labels to serve as axes for the analysis. One can still talk of "3rd person" in a purely semantic, referential, signifié sense, "neither speaker nor addressee", and use "first person" as, prototypically, a synonym for "speaker"--though notoriously one runs into paradox with "first person plural", since with that, one has abandoned the purely semantic and is making a grammatical claim--what claim is unclear. But as classificatory features they come with the analysis:

kuh: 1 sg	--	mih: 1 pl exclusive
--	tah: 1 dual, incl.	tam: 1 pl inclusive
muh: 2 sg	--	yuh: 2 pl
yah: 3 sg	--	dah: 3 pl

This analysis he finds "hardly elegant, economical, or convincing". He adduces no concrete objections; I assume that the three adjectives all point to the same fact: lacunae. The analysis he proposes

is "more satisfactory, economical, and semantically verifiable" and relies on a different set of features: \pm Speaker included in the reference, \pm Hearer included, \pm Minimal membership. This idea of 'semantic verifiability' is unclear: as simple definitions by extension, none of the labels "first person" (+ iff included speaker) etc. need verification; they're simply there, for better or worse. For Conklin, both verification and satisfactoriness seem to stem from a notion of economy. The analysis, then: dah -M, -S, -H; yuh -M, -S, +h, etc. A diagram makes clear that all the campers are in their bunks:



No reason is given for the excellence of this analysis beyond the fact that every combination is realized and the features are independent. And quite possibly this is the best analysis; but I suggest that one has as yet only one, very abstract and untested formal criterion to believe it. It may further actually mislead by suggesting that the addition or loss of a pronoun would be a disruption of a perfectly balanced system, comparable to the shattering of a crystal, the defacement of a work of art; whereas the asymmetrical array suggests that either tah may become lost or the language might want to forge another dual or so. Indeed, Germanic seems to have done just that, cf. OE wit, zit, only to discard them again. And in fact, overhauls of pronoun 'systems' are extremely common. Further, since pronouns are extremely frequent items of the vocabulary and tend to be short and pronounceable, we cannot plead the sort of erosion and non-systemic variation found in [k-/ik-ənamlks], haplogy, methatesis, [æpocowp], anapatictic, etc. Where there is System, there is Systemzwang: yet this one would almost seem to be self-sublating.

Conklin's formalism is embraced with a vengeance by Buchler and Freeze in an article in Anthropological Linguistics (Nov. 1966). They plead: "Although our approach may seem excessively formal and devoid of sociological relevance, it should be made clear at the outset that our goal is to analyze (through a uniform 'scoring procedure') a relatively large number of pronominal systems in order to isolate a set of components that will form an 'etic core' for future accounts which stress functional unity rather than dimensional range within a specific domain." This goal will appeal to all who feel that linguistics isn't a science until it's stored on a computer at Stanford, but the statement illustrates a sort of fallacy of composition: as though a piling-up of undermotivated and stillborn analyses would yield a sensuous total picture. An 'emic' study, we are informed, is one which is immanent to the language under investigation, while an 'etic' one is objective and

universal. But while a machine can be programmed to discover recurrent energy-distributions corresponding to (humanly pre-selected) features like "Voiced" or "Grave", no machine yields Buchler and Freeze's would-be objective (transcendent, ahistorical) features as the relevant axes for formal analysis, such as \pm Min, \pm Solidarity. Aware of the problem, the authors say that "the distinction between emics, on the one hand, and etics, on the other, should be formulated in terms of different levels of structural validity, rather than in terms of a contrast between psychologically 'real' and structurally 'real' descriptions." This looks promising, but no hint is given of the content of any of these notions, particularly when the analysis concerns an exotic language which they are not otherwise describing for us. I shall try to give some content to the notion 'structural validity' as regards pronoun features, and will suggest that little content accrues to it if we restrict our view to a set of half a dozen or so pronouns together with their glosses.

To see that this is no niggling criticism, consider their analysis of Totonac person morphemes: acquit "I"; aquin "we"; ama "he, she, they"; huix "you (sg.)"; huixin "you (≥ 2)". [Space prevents inclusion of discussion of their analysis in terms of secondary features Min, Max, Spkr. The upshot is: it is a mere parlor trick.] One would have to look at more of the language to know whether this analysis had any advantage over any other analysis. Nothing has been added by the labels; such meat as there is lies in the raw glosses, which represent not analysis but reportage. Thus, Totonac suppresses the sg/pl distinction in the third person--to use the traditional labels--English in the second, so for both one could propose this kind of analysis: in English you, in Totonac ama receive the lone + value for Max. Now we know that this would be off-the-wall for English (our intuition on this score is composed of our knowledge of the facts of verb agreement, etc.), so why should it be appropriate for Totonac? It might be; we simply have seen no evidence. Moreover, what little we glimpse here of the morphology suggests that first person (aqui-) and second person (huix-) are not without expression in Totonac, as is plural (-in).

A caution about 'etic' analysis! I doubt that as an absolute type there is any such thing: a (necessarily partial) reproduction of the data is more or less detailed, more or less influenced by considerations of contrast, more or less reshuffled. I'd be willing for sake or argument to admit the list of glosses for the pronouns as fairly 'etic', fairly 'given on a platter by the facts', but not the controversial further features. And, of course, even the glosses are idealized, 'emicized' almost, since no doubt every language allows its pronouns to stray in ways the analyst cannot account for a priori. Thus he but not she is used for anaphora of one in American English (but usually one again in British English) or as anaphor of distributively understood everyone, anyone; there is capricious variance between it and he referring to animals (further: cat, she; dog, he), etc. All that we toss out, because

we have identified the core of these pronouns--not an etic notion, this core--on a broad consideration of pattern. Similarly, even the 'most narrowly' phonetic of transcriptions (i.e., not "the narrowest conceivable", which is an undefinable notion, but "the narrowest we in practice come across") picks and chooses among the data it is to relate ...

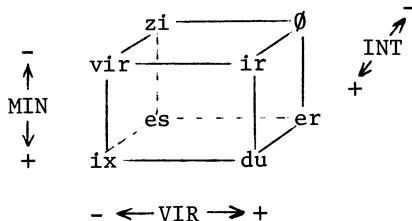
What of our two abstract conditions for the satisfactoriness of analyses--binary and independent distinctive features? It is hard to see what justification they have other than operational: striving towards such an analysis at least keeps one from lazy ones such as, in a limiting case, having as many features as elements, say "Egoism" for ich, all the other pronouns being -Egoism; "Dutztum" for du, etc., in which case one's n elements are at equidistant corners of a very slimly filled n-dimensional tesseract; or at the other extreme, a single n-ary feature "Ubu", with values U_1 a.k.a. Egoism at ich, U_2 at du ... in which case we are back to a set with no structure. One could be convinced of, say, the validity of the binary principle in phonology, but as the result of grueling empirical study packed with all sorts of surprises, mostly peculiar to phonology. Cf. Whorf: "We do well to be skeptical of a grammarian's systematization when it is full of ENANTIOMORPHISM, the pairing with every category of an opposite which is merely the lack of it." Martinet suggests (Ling. Synch., p. 88) that 'binarisme' is simply the result of the practical necessity of comparing elements pairwise.

We must, then, look at more of the language than the tiny subset Y to determine whether our componential analysis of Y has even a claim to being valid. It is no defense to claim that you are merely doing an etic analysis, if the sense of this term is not to fall to the level of "random selection from the kaleidoscopic set of combinatorial possibilities." (For a census of this set, see R. Burling, American Anthropologist, 1964.)

Take English. Our gloss for you (apart from cases where it doesn't refer at all, the 'impersonal' you) indicates that it may refer to any number of people; and because you (sg) comes from you (pl) historically and has kept the verbal concord of the latter, the Sg/Pl distinction (or status of the feature \pm Pl) isn't so neat. So far as this goes, we might say that the second person is unmarked for number and talk rather of 'contextual determination' à la Buchler and Freeze. But if we look at the reflexive forms, our distinction reappears: myself:ourselves::yourself:yourselves. This is evidence for the relevance of our feature \pm Pl for pronouns, even if the distinction is occasionally suppressed in some of them. This is especially true as the sg:pl in -self:-selves is evidently the same as the sg:pl in the homonymous independent nouns. The occurrence of the feature in the class of Nouns is our main justification for using the same label for a like distinction in the class of Pronouns; or better, its relatively regular expression in the open class of Nouns is the foundation of its (relatively defective) application in the paltry class of Pronouns.

The case of German may reveal some of the sinews and arteries

concealed beneath the integument of the distinctive-feature cube. To start with, we have only unordered items with glosses: ix "I"; vir "we"; er "he"; es "it"; du "you; spoken to a single referent with whom the speaker is, roughly, intimate"; ir "you plural, intimate"; zi "she; they; you singular/plural nonintimate." At this point our atomist can sort things out in terms of, say, Minimality (+ if the expression can never refer to more than one), Intimacy (+ if someone in the referent is intimate with the speaker, who is of course intimate with himself), and 'Virility' (necessarily refers to masculine if speaker is feminine and all addressees male):



If the speakers belong to an exotic tribe (say the Alemanni), one can get away with such things, and even--having arrived at the 'analysis' via aprioristic linguistics--relate the analysis to social facts ("The males mark Virility by means of distinctive clothing variables."). We know this won't do here, because we know German. But how much German do we have to know, to know this? And is even German enough? That is, in general, how widely must we cast our net, to catch a given fish? [Omitted: discussion of Chomsky's position that, while other levels may be there to be described, if reference to them is not necessary then you are breaking the rules of the game.] In roaming far afield for justification of one's analysis of a small subset, one always risks missing a key closer to home. But surely it is quite as common for formalists not exactly to miss, but to decline to use a common-sense (because functional) key, circumventing it with elaborate contortions, out of respect for the separation of levels. Z. Harris, a heroically consistent researcher in this line, is aware that his analyses tend to be "laboriously distributional", but counters that "since there is no independently known structure of meanings which exactly parallels linguistic structure, we cannot mix distributional investigations with occasional assists from meaning whenever the going is hard". As a result, watching Harris do a discourse analysis of a text of English as though it were Linear A, referring only to morpheme boundaries, is like attending an improvisation for children, in which the actor pretends not to notice the rabbit (or pirate) in full view behind his back, and persists in his elaborately ineffectual peregrinations, despite howls and gesticulations from the audience.

Let us try to arrange the German data in accordance with our already existing intuitions, as reflected in (or possibly merely engendered by) the traditional labeling, and see what kind of facts one must adduce in support.

The first scission may be made on the basis of the Pronoun set itself, if we include oblique forms: for /zi/ "she" becomes /ir/ in the dative, while /zi/ "you, they" becomes /inən/. But this is not overwhelming evidence; it does not force us to conclude that sie and Sie (orthographic forms respectively for "she" and "you"; "they" again awkwardly sie; I shall write SIE for the union, which we have yet to split other than referentially and orthographically, of Sie and sie (pl)) are as separate as ich + du any more than we are forced to conclude that there are two unrelated lexemes /vɔrt/ ("word"), one with plural /vɔrtə/ "paroles", one with /vɔrtər/ "mots". After all, sie and SIE share accusative, genitive, and even the dative, which suggests the same base morpheme: zi (nom. sg. fem.) ir (dat. sg.):zi (nom. pl.):inən (dat. pl.) is close to the deictic/definite article: di (nom. sg. fem.):dər (dat. sg.):di (nom. pl.):denən (dat. pl.), allowance made for the fuzzier morphology one generally finds in Pronouns. No, to decide the case we must go outside the roster of Pronouns proper. Let us take the smallest step, to syntactic constructions involving Pronouns. Relief in the "sie/SIE" confusion is immediate: sie takes ist, luft, etc.; SIE takes sind, laufen. This provides a formal method of distinguishing the two pronouns ("two" because SIE at this point must still be considered a single Pronoun, albeit with a curiously broad gloss). But what sort of distinction is this--what, beyond saying that A ≠ B? We cannot say without looking at still more of the language. So we consider constructions that do not even contain pronouns, like "Die Frau ist alt/Die Frauen sind alt", and discover that the Pronoun - Verb concord parallels what for nouns, on overwhelming evidence, we would want to call a Singular vs. Plural distinction. This is, of course, the reason for calling (traditionally) wir "first person plural", despite the semantic nonsense of the label (WE are not simply a bunch of ME's). Thus, although we have started from a notional category (first person = the one who is speaking), we have implicitly yielded to the primacy of the formal (surface concord, morphology) in establishing our technical labels.

A primary justification for the use of features in phonology, or of word-classes in syntax, is that segments sharing these features are treated in the same way by various rules (and since the domain of these rules is unlimited, this has predictive value), a fact sometimes expressed by saying that the rules mention the features, thus emphasizing the status of these features as (among other things) labels. Let us see how the time-tested features of first, second, third person, singular and plural for German Pronouns stand up under this criterion. We have already seen that Singular: Plural holds up well in third person. The rest will not come so easily.

Almost none of the rules I shall mention distinguish er, es, sie (sg.), so I shall lump these for the time being as ER: they will have to be distinguished later by another feature, Gender. This distinction, like that of Number, is based on the close analogy with a feature so labeled and much more easily established (adjectival paradigms) for Nouns. Our rules also fail to distinguish, save

marginally, between Sie (singular, plural) and sie (plural), which we have therefore lumped as SIE; the unpacking of this one will be more complicated. For now, call it third person. In what follows, we circle those terms which are treated identically by a given rule.

First, our third person is treated uniformly by a few rules--rules logically independent of one another, which strengthens our sense of the reality of the category; e.g.

ich	wir	(1)	reflexive Pronoun = <u>sich</u>
du	ihr	(α)	(2) lack of obligatory Pronoun copy in relative clause: e.g., Du, der du noch ein Kind bist; Max, der (*er)...; and even, anticipating a bit, Jesu [voc], der du meine Seele hast durch deinen bitteren Tod...
<u>ER</u>	<u>SIE</u>		

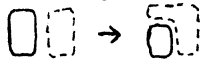
Next, Singular is distinguished:

<u>ich</u>	wir		<u>darf-</u> base in Present
<u>du</u>	ihr	(β)	<u>muss-</u> base in Present
<u>ER</u>	SIE		similarly <u>mag-</u> , <u>kann-</u> , <u>will-</u>

(Note that semantically singular Sie "Usted" is Plural in this rule.) Singular was also distinguished by a wider-ranging rule of Middle High German: e, ʏ, ʊ → i, ie in the stem of the present of certain verbs (sprechen, gebären...). But today we see rather this pattern:

ich	wir	(1)	e, ʏ, ʊ → i, ie in certain Verbs (<u>sprechen...</u>)
<u>du</u>	ihr	(γ)	(2) a, o, au → ʌ, ʝ, ʒu in certain Verbs (<u>laufen...</u>)
<u>ER</u>	SIE		

The change has been said to occur by 'analogy'.

But this is awkward. For if ich, du, ER share a feature which ich, wir, ihr, SIE do not, in this case Singular, and if the converse does not hold, then we have here not analogical leveling but, seemingly, capricious terracing:  We may attempt to buttress the Number distinction by the facts of sein concord, in the Present:

<u>ich</u>	<u>wir</u>	---	<u>sind</u>
<u>du</u>	<u>ihr</u>	(δ)	--- apparently shared <u>bi-</u> base
<u>ER</u>	<u>SIE</u>		... apparently shared <u>sV-</u> base

But this, like that adduced previously, is weak and idiosyncratic evidence for the overall structural validity of the traditional distinctions (and is even historically ill-founded in the IE *es- appears in ER as well as wir, ihr, SIE verb-forms). Vastly more widespread, comprising an indefinite number of Verbs, is the following curious pattern, observed by rules of which I list a handful:

<u>ich</u>	<u>wir</u>	(ε)	(1) Past concord with most weak verbs, of various subclasses-- <u>loben</u> , <u>reden</u>
<u>du</u>	<u>ihr</u>		
<u>ER</u>	<u>SIE</u>		

(2) Konjunktiv I, II of various strong and weak verbs (dürfen, laufen, brachen, binden...)

(3) sein concord in Past Konjunktiv I, II. (4) modals; various strong preterites.

Viewed one way, this is strong and distressing evidence against the linguistic reality of the traditional Pronoun categories of either Number or Person. Viewed another, it is grudging confirmation of these categories. And our eticist brethren have no basis on which to choose: they have no basis on which to proceed past the original unstructured list of glosses.

The view I have in mind would not take (ε) at 'face value' (i.e. would not consider that it established categories ich + ER, wir + SIE) but would consider the apparent pattern to be a superposition of second person marking upon a straight singular vs. plural distinction. But how are we to justify this sleight of hand?

Since the point of this paper is not to establish this or that particular analysis of pronouns, but to examine the sort of evidence that must be brought to bear in evaluating the appropriateness of grammatical labels and categories (this in turn being relevant to the characterization of their ontological status), I'll not attempt to adduce every shred of support that one might scrounge, but simply point out that to decide or even examine the matter we must open our court to yet another class of witness. To set up (ε), we had to look at a lot of the language beyond the roster of pronouns, but we did it in a way that was both surfacey and local: a computer, suitably instructed about morphology, could have drawn it on the basis of distributional facts, without holding the entire language in its regard at one time. But now it is suggested that (ε) is in 'fact' a superposition of two other patterns, neither of which perhaps overtly occurs! One recalls disputes as to the correct Deep Structure for a sentence, all candidates equally and elusively in the nether world.

One can make a general plausibility-argument for why the second person is strongly enough marked that it could cut into a conjugational pattern to the extent of disrupting a sharp Singular/Plural distinction. The argument is subjective because it is not clear what weight to assign the various sorts of evidence. (1) The second person alone is deemed worthy by the language of an Intimate-Nonintimate distinction--even if the nonintimate semantically second person forms were pilfered from the formally third person. Thus, in a completely different sphere, the second person is sharply characterized. [Rest of discussion omitted.] Whatever one decides with regard to (ε), there is another pattern of open application for which no principled interpretation exists: Present concord with weak Verbs of several subclasses and non-umlauting strong Verbs (loben; betrügen):

ich wir
 du ihr
ER SIE

Or again, present of umlauting strong Verbs and most modals:

ich wir
 du ihr
 ER SIE

It is such jigsaw pieces, and not any pristine system of uniform

oppositions, that determines the topography of the personal pronouns. We've had to do a fair amount of empirical mucking-about just to establish the roster of personal pronouns, let alone their relative interdistances.

What are we to make of our uncomfortable result--so far that Sie "you" is not distinguished from its third person homophone? [Discussion omitted. Upshot: There is indeed little formal distinction--"Es bleibt ein Fleck im Gewand der deutschen Sprache, den wir nicht mehr auswaschen können." (Grimm) If we do, as speakers, feel the two plural-concord Sie's to be quite distinct, then the formal is here bowing before the social.] We illustrate the rule that does separate Sie "usted" from its two semantically-plural homophones, for the record:

ich	wir	Collocable with <u>einander</u> .
du	ihr	
Sie	Sie	It is, as we remarked, cryptosemantic.
ER	sie (pl)	

We next note a little structure in the ER group:

ich	wir	---	<u>ihr</u> -	possessive,	<u>sie</u>	nominative	accusative
du	ihr	---	<u>sein</u> -	possessive,	<u>ihm</u>	dative	
<u>SIE</u>	<u>Sie</u>						

The distinction we want is much like the well-founded distinction of gender in nouns, so we may use the same labels for the Pronouns.

Current usage presents a blurred picture--or rather a moiré, owing to the competition between two systems of pronominal reference, formal and semantic (grammatical and natural). In thus admitting Gender (Masculine, Feminine, Neuter) as a dimension of Pronoun classification, we have, however, beyond offending the binarophiles, offended the partisans of symmetry, who note massive syncretism in the resultant three-dimensional (so far: Number, Person, Gender) entabulation, and correctly point out that, if we restrict our gaze to the Pronoun set, our establishing a dimension to handle just one of our six major pronoun-bundles looks like a case of special pleading. And indeed, to justify our metaphor of 'Gender' for Pronouns we must look, not just at the Pronoun set, nor even just at German, but at languages generally: our widest scope so far. We decide, after looking at lots of languages (or, as a short-course, reading J. Greenberg "Language Universals") that Plural is a marked category, that first and second person are marked with respect to third, and that distinctions made in unmarked categories tend to be felt as an overload in marked ones: hence the particular syncretisms we have to confront here are not a refutation of the system. That is, there are gaps and there are gaps, and these are lawful.

That there are gaps is, however, not without significance, and perhaps suggests a rather low functional load to Gender in modern German (a suggestion that would not arise if we tried a priori to stuff all 'systems' into relative economy): thus we might not be surprised to learn that it has further retreated in other Germanic

languages: marginal in English Nouns and often semantically non-minimal (steward, stewardess), and reduced to two in Scandinavian, except in Old Icelandic, where, however, all three genders were marked in the plural as well.

One thing we might like from a description of a system, something akin to explanation, is an insight into the historical development of that system, in cases where the material under description is empirically unstable. Does our net of linguistic features tell us anything about the changing constitution of the German Pronoun 'system' over the past few centuries? Unfortunately I cannot see that it does. The motivating force behind the extensive upheavals in the Pronoun system seems to have been the jockeying for position among the castes and classes, rather than any inherent linguistic tension in the MHG system or inherent advantage of the present system: if anything, a previously well-balanced system has been battered repeatedly, and what we have left is a sort of rubble, full of ambiguities and asymmetries. There have been deletions, additions, and rearrangements in the roster of the Pronouns themselves, including periods when there really wasn't any nominative second person Pronoun for certain addressees, and one had to make delicate choices among various more or less conventional NPs used quasi-vocatively. The hinge to all this self-abolition and renewal finds only the palest one-dimensional reflection in what is linguistically, in fact, the least well-grounded of our features, \pm Polite. The sprachliches Feld of German Pronouns is stained with the bile of spurned courtiers, with the tears of children chastized for addressing the grandfather with the form that the father is wont to receive, with the blood of students who, away from home and patois, took affront at Er from a townsman (Mainz and Göttingen, 18th century). In sum: pattern-maximizing treatment of an abstractly delimited subset is otiose; subjective, grammatically wide-ranging considerations are necessary for even a modest taxonomy, though in so doing one has come up with an interesting description; and explanation of anything genuinely interesting requires non-linguistic supplement. In general, to all abstract or purely synchronic linguistics, I would oppose the slogan: You can't take the class struggle out of language.