

Linguistic Economy And The Typology of Turkic Pronouns

Grace Mrowicki
Senior Thesis
Haverford College
Swarthmore College Department of Linguistics
December 18, 2003

Chapter 1: Introduction

Personal pronouns are a feature of language so basic and omnipresent that we rarely think about them. Perhaps it is because of their ubiquitous presence that they are, by nature, conceptually somewhat slippery. These bundles of meaning express, at the very least, the relationship to event (person; i.e. speaker, hearer, referent), but also may account for notions of number, gender, formality, animacy, and also for more esoteric classifications. However, over time, a given language may incorporate a new classification into its system of pronouns, or, often in the wake of incorporating a new classification, rework the system from the inside. It is this second type of change that this paper is particularly concerned with. Pronouns usually change in these situations to maintain distinction within the system. However, this type of releveling is, like all morphological change, subject to the constraint of memorability, although this constraint of memorability can be the impetus for such change. This is to say that pronouns should, theoretically, change in accordance with what is easiest to remember.

So what makes something easy to remember? That's where it gets fuzzy. What people can agree about is that language wants more for its money so to speak: it wants the most meaning with the least effort. There's a reason more phonological change can be attributed to lenition than fortition: lenition, by definition, is easier, and language wants easier. Most linguists will agree that the simplest possible morphological system involves a consistent one-to-one relationship of meaning to form. To use terms which will I will define more formally later, this idealized system is productive and transparent, all of which refer to the ability to parse out the functional pieces of a word and semantic

morphology, deal with the dynamic between accessibility (defined by brevity and/or morphological consistency) and distinctness.

Pronouns occupy a special place in terms of morphology since they are, by and large, composed of meanings commonly expressed in inflectional morphology, such as person and number. This is part of their susceptibility in regards to leveling. As an introduction to how the tug-of-war between accessibility of forms and distinctness of forms plays out in the pronominal paradigms of other languages, consider the trajectory of the English language, even just since the beginning of “Modern” English in the 14th century.

1) EARLY MODERN	singular	plural
1 st	I	we
2 nd familiar	thou	ye
2 nd formal	ye	
3 rd m./f./n.	he/she ¹ /it	they

2) CURRENT STANDARD	singular	plural
1 st	I	we
2 nd	you	you
3 rd m./f./n.	he/she/it	they

A note about shading:

The charts of pronouns are shaded to show where a form occurs multiple places as a result of using one form for multiple permutations of information (i.e. using an existing plural form for formal address of one person).

Modern Standard English has lost any distinction of formality or number in the 2nd person. In fact, it is these concepts of formality and number, especially in the 2nd person, that seem to be the most volatile in language. Consider the following few examples from relatively familiar languages:

¹ I star this because *she* had not fully codified. Howe (1996) notes that also in use for the 3rd person feminine at this time were *scho*, *heo*, *hu*, *ha*, *he*, in different dialect regions of England

5) SPANISH	singular	plural
1 st	yo	nosotros
2 nd familiar	tú	vosotros
2 nd formal	usted	ustedes
3 rd m./f./n.	él/ella/ello	ellos/ellas

6) GERMAN	singular	plural
1 st	ich	wir
2 nd familiar	du	ihr
2 nd formal	Sie	Sie
3 rd m./f./n.	er/sie/es	sie

In Spanish the 2nd person formal *usted* is a reduced version of *vuestra merced* (your mercy), and behaves grammatically 3rd person singular pronoun. In the German example, they take advantage once again of their creative use of capitalization, and *Sie* behaves grammatically like the 3rd person plural pronoun *sie*.⁴

However, using existing pronominal forms to create a T/V distinction causes a loss of distinction elsewhere in the paradigm. In the case of a language like Russian, using *vy* to indicate 2nd sg formal means that in the 2nd sg, there is a clear distinction between familiar and formal. However, it means that you can't always grammatically distinguish 2nd pl from 2nd formal, since they are both *vy*. The German *sie* and *Sie*, since they are phonologically indistinguishable, also create a loss of distinction in the paradigm. And nowhere is this more true than in English, where the loss of the 2nd sg. fam. *thou*, in combination with the use of the 2nd pl *you* to indicate 2nd formal, means that in Modern Standard English, *you* covers the entirety of the 2nd person, and there are no distinctions for number or formality anymore.

However, English speakers have been improvising ways to redistinguish plurality in the 2nd person paradigm. Commonly, this is done by adding words that indicate a

⁴ Distinguished from 3rd sg fem *sie* because of dative form: 3rd pl/2nd form dat *ihnen*; 3rd sg fem dat *ihr*.

only grammatically accounts for person and number, so there are only two meaningful components of these pronouns. It appears to break down as follows:

- [b] corresponds to 1st person
- [s] corresponds to 2nd person
- [en] corresponds to singular
- [iz] corresponds to plural

Formality is indicated through the familiar model of using the 2nd pl for this function.

This system seems pretty basic and easy to comprehend, much in line with the discussion of an idealized one-to-one meaning to form type system. There appears to be an inherent logic to what parts mean which, as opposed to the whole thing being a feature bundle, as in the case of most of the Indo-European examples.

However, looking at Kazak, a Western Turkic language that the Turkish falsely regard as a dialect of their own language, the picture is similar but slightly more complicated:

10) KAZAK	Singular	Plural
1 st	men	biz
2 nd familiar	sen	sender
2 nd formal	siz	sizder
3 rd	ol	olar

Here we see pretty much all of the forms from the Turkish paradigm assuming that the [m] in the 1st sg here can be explained as an allophone with the [b] that turns up in all the other Turkic 1st person examples presented thus far of an underlying /+labial/⁶. However, there are also distinct forms for each of the eight permutations of information. The 2nd pl forms appear structurally a bit more complex, with the addition of [der], a form of the pan-Turkic pluralizing suffix /LAr/ on top of the forms with /iz/, the apparent pluralizer

⁶ Turkic languages are full of allophony, more of which will turn up later, based on assimilation of voicing/continuity along the same place of articulation, and 1st person morphology shows a lot of allophony among the labial consonants.

schema of Turkish (which generally follows an idealized one-to-one relationship) or of Kazak (which deviates from this somehow by incorporating /LAr/ into the 2nd person pl).⁷

Because of these factors and growing inconsistencies, Turkic pronouns provide an interesting perspective on the phenomenon of 2nd person formality and pronoun change in respect to theories of memorability because of the explicitness of Turkic agglutinating morphology and the morphological conflicts that arise from the fairly common processes involved. The other interesting point surrounding this change in Turkic pronominal morphology is its place in the argument that agglutinating morphology tends, over time, to evolve into fusional morphology, as has happened in Estonian and most Lappish dialects. Most linguists who make this claim attribute this to a collapse of transparent morphological boundaries due to phonological change. However, what by and large is seen in these examples is a loss of perception of transparency without a phonological impetus.

In various grammars, which is to say in the context of individual languages, authors have offhandedly tried to explain /iz/. On the other hand, Swift's *Reference Grammar of Modern Turkish* (1963) only discusses /iz/ as a personal pluralizer suffix. Since Turkish is one of the languages for which /iz/ consistently works as a pluralizing suffix in the 1st and 2nd person, it's not an issue for discussion much in that context and as such this analysis of /iz/ for Turkish makes perfect sense. von Gabain's *Alttürkische Grammatik* (Old Turkic Grammar) (1950) describes this (I)z morpheme as a no longer productive dual suffix used in cases of natural pairs such as eyes (Gen. Turkic: *köz*).

⁷ see Appendix I for a summary of the pronoun paradigms for as many Turkic languages as I could get my hands on

of their distribution. In the 1st person here, there are separate, phonologically unrelated morphemes for singular and plural.

I have referred to these sets of pronouns as paradigms. What defines a paradigm, and what defines these sets of pronouns as a paradigm? Carstairs (1987) provides a good definition of paradigm as follows:

A paradigm for a part of speech N in a language L is a pattern P of inflectional realizations for all combinations of non-lexically-determined morphosyntactic properties associated with N such that some member of N exemplifies P (i.e. displays all and only realizations in P). (Carstairs 1987: 48-9)

This definition builds its paradigms from the data upwards, as opposed to from a feature of a word downwards. This allows for examples like the Latin nouns *dominus* (masc.) and *fraxinus* (fem.) to be part of the same paradigm, since they follow the same inflectional pattern rather than referring to their gender, which would predict distinct sets of inflectional exponents. Also useful is Wurzel's (1984: 66) definition of inflectional class, a notion quite similar to that of paradigm, such that if two words express the same morphosyntactic properties by phonologically unrelatable exponents, they belong to different inflectional classes.

Carstairs (1987) is interested in the phenomenon of allomorphy in inflectional morphology. This is to say, different ways of indicating a given morphosyntactic property within a language. He describes four different ways that inflectional morphology deviates from the idealized one-to-one relationship of morphosyntactic property to inflectional realization:

- I. One property to many exponents syntagmatically
("double/multiple-marking" more than one morph is used to express a single notion in one word)
- II. One property to many exponents paradigmatically

example of person and number agreement, where 3 persons are recognized by morphemes a, b, and c and plural is recognized by morpheme p, a speaker has to memorize the following:

Table 2.3

agglutinating model	sg	pl
1	a	a p
2	b	b p
3	c	c p

In order to produce these 6 agreement markers, the speaker needs only remember 4 pieces of information: a, b, c, and p. Now let's look at a model which involves homonymy

Table 2.4

agglutinating w/homonymy	sg	pl
1	a	a p
2	b	b p
3	c	b p

In 2.2, the speaker has the same 4 morphemes to remember as in Table 2.1, but also must remember that b is used for the 3rd person in the plural instead of c, so that rule is counted as the 5th thing that must be remembered. This means that the model in Table 2.2 has more to remember than that in Table 2.1, is therefore more work to memorize and is therefore less likely to survive.

What does it mean for morphology to be transparent? Earlier I mentioned the connection between transparency and productivity in that they both have to do with the ability to parse out the different morphological components of a word. Mayerthaler (1981) puts together a good formal definition of transparency which describes a four-part gradation of transparency. He bases his definition on Brame's (1974) 'natural bracketing

- (d) a word is ‘constructionally opaque’ when (1) and/or (2) are not applicable and the word does not appear to have any construction seams
 e.g. book
 (Mayerthaler 1981:96, examples mine)

With these formal definitions in mind, let’s look at how Turkic pronouns measure up here:

Table 2.5

TURKISH PRONOUNS	Singular	Plural
1 st	b-en	b-iz
2 nd	s-en	s-iz
3 rd	on	on-lar

- [b] corresponds to 1st person
 [s] corresponds to 2nd person
 [en] corresponds to singular
 [iz] corresponds to plural

Turkish, unsurprisingly, breaks down along Mayerthaler’s criteria for type (a) full transparency, as observed in Chapter 1, clearly into person and number components. The parsing out of /iz/ in this is useful because of its application to other person ‘roots’ in affixal forms.

Kazak is not so clear:

Table 2.6

KAZAK PRONOUNS	Singular	Plural
1 st	men	biz
2 nd familiar	sen	sender
2 nd formal	siz	sizder
3 rd	ol	olar

Generalizations for whole paradigm:

- (a) /+labial/ corresponds to 1st person
 (b) /s/ corresponds to 2nd person

1st person generalizations:

- (c) /en/ corresponds to singular
 (d) /iz/ corresponds to plural

The second possible analysis of these types of pronouns is that /iz/ has ceased to be productive. By this analysis, the model breaks down as follows (representations: men=a, sen=b, siz=c, ol=d, biz=e; /ler/=1)

Table 2.8

Kazak model 2	sg	pl
1 st	a	e
2 nd familiar	b	b 1
2 nd formal	c	c 1
3 rd	d	d 1

This model requires the speaker to memorize 6 morphemes and one rule (that 1pl does not get marked by /ler/), for a total of 7 things to memorize for 8 forms. This is also the way that the affixal paradigm in Table 2.1 breaks down. As compared to the other analysis, this is quantitatively easier to memorize. It also reclassifies the way that it deviates from the one-to-one model. This analysis goes along with Deviation III, the one involving ascribing more than one feature to one morph. This is also the deviation that was used to describe the fusional Indo-European examples.

This simpler model is counterintuitive because of the seeming transparency of it. It corresponds to Mayerthaler's definition of non-transparency, which is to say that it looks like it should break down farther than it actually does. Taking the persons separately, they seem obviously parsable, and I spoke earlier to the usefulness of parsing out /iz/ since it is attached to other 'person' roots in postlexical agreement forms, where this system does not allow that. However, there's another interesting piece of data that lends itself to the non-transparent interpretation of these pronouns. Relatively recently, scholars of Tuvan have reported that the form *silerler* is being attested as a 2nd pl formal, in contrast with the usage of *siler* as 2nd sg formal (Anderson and Harrison, 1999). While it would seem that [ler] as a regular form of /LAr/ would be remarkably transparent as

Chapter 3: What is to become of *us*?

However, this focus on plural marking of the 2nd person leaves the case of the 1st pl hanging. The rule that must be remembered in the example of Kazak-type paradigms is that the 1st pl does not need nor take /LAr/ to indicate that it's plural. Part of this is accounted for by the fact that because it uses an independent form from the 1st sg and that 1st pl form is not used for anything else, the notion of plurality is already accounted for in it; To tack on the /LAr/ at the end would actually double mark plurality here, as opposed to the 2nd person, where it creates a useful distinction. However, by the analysis I have been using, to add the /LAr/ would make the paradigm ostensibly easier to remember, since you wouldn't have to remember the rule which said that the 1st pl does not take /LAr/.

Carstairs theory remains curiously silent on the subject of redundancy of information, aside from examples given in regards to Deviation I, which describes one property to many exponents in the context of one example, but none of his examples included using the same kind of affix more than once. It's also notable here that [is], derived historically from /iz/ is only fully present in the 1st pl form. There is *sen* and *siler*. Of other Turkic languages not extensively discussed, Uighur and Kyrgyz also display *siler* forms, though both of them also have *siz*, and as such neither has borrowed *siler* into the formal singular.

A few examples may be of use here: unrelated Basque, and a few Turkic languages less discussed here, for which I have seen incomplete data that includes use of /LAr/ in the 1st person plural (generally speaking, *biz*+/LAr/). These languages include Tuvan, Uzbek (361), Altai (409), Noghai (336), Crimean Tatar (307), Turkish (209),

Description of these 1st pl forms involving /LAr/ as individualizing is inconsistent, but the seeming lack of influence on agreement morphology is consistent, or at least the lack of comment on it is. None of the data on any of the languages where 1st pl + /LAr/ forms have been documented seem to demonstrate any difference in pronominal agreement morphology, which, as established earlier, more or less follows the same morphological patterns as the personal pronouns. This inconsistency leads me to believe that either people are ignoring it, or that they behave similarly and as such might be the same phenomenon, one way or the other. It is quite possible that some of the people listing these forms haven't thought they might be individualizing plurals, more likely I think, than the other way around.

Basque, while genetically unrelated, is interesting because its pronoun paradigm displays a strikingly similar pattern to the Kazak-type Turkic paradigms, which suggests a very similar process, and is also generally considered to have an agglutinating-type morphology.

Table 3.1

BASQUE PRONOUNS	singular	plural
1	ni	gu
2 familiar	hi	zuek
2 formal	zu	
3 ⁸	bera	berak

While Basque not quite as frustrating in its non-transparency as the Turkic patterns, since there is no real phonological association between [n] and [g] or between [h] and [z], to create any real semblance of transparency, the pattern of the final vowels is notable, with an [i] on the 1st and 2nd familiar singular (think /en/ in the Turkic examples), [u] on the 1st

⁸ Basque has not historically used 3rd person pronouns. These forms are derived from 3rd person emphatics, and are not in use in all dialects of Basque. (Trask 1996)

most meaning with the least effort. So how does redundancy manifest here in a way that increases effort, but not meaning? What makes *silerler* acceptable, but **bister* not? Both of these forms add general plural marking onto what appears to be a plurally marked form. However, as discussed earlier, because *siler* is derived from *sen*, but *sen* has been reduced to *si-*, it's easier to accept the *-ler* as part of the whole form than as a regularly affixed form of /LAr/. Because it is then in use as the 2nd sg formal, in addition to 2nd general/familiar pl, it is not redundant to tack on /LAr/. As long as we accept *siler* as being not inherently plural, the addition of /LAr/ gives more information. **Bister*, on the other hand, does not provide any new information. It is purely superfluous. Of course, another consideration in regards to the non-proliferation of /LAr/ in the 1st pl would be another economy constraint. It ostensibly takes more effort to say the extra syllable, and if it's purposeless, people are going to be even less inclined to want to say it.

Carstairs' Deviation I (one-to-many syntagmatically), which seems to be the only thing I can find on multiple marking of features, is distinct from this. The examples Carstairs cites seem to be examples of what I would call partial redundancy: these examples deal with instances of stem change and/or fusional morphology that repeat some of the information, but provide additional information wrapped up with it, and so are not analogous instances of redundancy.

Appendix 1: Turkic Pronouns

Table A: Pronoun chart marked for actual usage (light gray = sg. formal; black = plural; dark gray = pl. formal)

	Old Turkic	Kazak	Kyrgyz	Turkish	Bashkir	Uzbek	Uighur	Tuvan
1 sg	m(ä)n	men	men	ben	min	men	män	men
2 sg fam	s(ä)n	sen	sen	sen	hin	sen	sän	sen
2 sg form	siz	siz	siz	siz	heð	siz	siz	siler
3 sg	ol	ol	al	on	ul	u	u	ol
1 pl	biz	biz	biz	biz	beð	biz	biz	bis(ter)
2 pl fam	siz	sender	siler	siz	heð	siz	silä(r)	siler
2 pl form		sizder	sizder					silerler
3 pl	olar	olar	alar	onlar	ular	ular	ular	olar

TABLE B: Pronoun Chart Marked for Presence of /iz/ and /LAr/ (light gray=/iz/; black=/LAr/; darker gray=both /iz/ and /LAr/; darkest gray= double use of /LAr/)

	Old Turkic	Kazak	Kyrgyz	Turkish	Bashkir	Uzbek	Uighur	Tuvan
1 sg	m(ä)n	men	men	ben	min	men	män	men
2 sg fam	s(ä)n	sen	sen	sen	hin	sen	sän	sen
2 sg form	siz	siz	siz	siz	heð	siz	siz	siler
3 sg	ol	ol	al	on	ul	u	u	ol
1 pl	biz	biz	biz	biz	beð	biz	biz	bis(ter)
2 pl fam	siz	sender	siler	siz	heð	siz		siler
2 pl form		sizder	sizder				silä(r)	silerler
3 pl	olar	olar	alar	onlar	ular	ular	ular	olar

Appendix II: Some examples of Turkic agreement morphology

Kazak Cliticized Pro. Agr.	Singular	Plural
1	-MIn	-MIz
2 familiar	-sIn	-sInAr
2 formal	-sIz	-sIzdAr
3	-dI	-dI

Uzbek Cliticized Pro. Agr.	Singular	Plural
1	-mān	-miz
2	-sān	-siz
3	null	-lār

Kazak Affixal Pro. Agr. 1	Singular	Plural
1	-m	-mIz
2 familiar	-ŋ	-ŋIz
2 formal	-ŋIz	-ŋIz
3	-I, -sI	-I, -sI

Uzbek Affixal Pro. Agr. 1	Singular	Plural
1 st	-m	-miz
2 nd	-ŋ	-ŋiz
3 rd	-i, -si	-i, -si

Kazak Affixal Pro. Agr. 2	Singular	Plural
1	-m	-K
2 familiar	-ŋ	-ŋdEr
2 formal	-ŋIz	-ŋIzdEr
3	null	null

Uzbek Affixal Pro Agr. 2	Singular	Plural
1	-m	-K
2	-ŋ	-ŋiz
3	null	null

Bibliography

- Swift, Lloyd B. *A Reference Grammar of Modern Turkish*. Bloomington, IN: Indiana University Publications, 1963.
- The Turkic Languages*, eds. Lars Johanson and Éva Á Csató. London: Routledge, 1998.
Éva Csató and Birsel Karakoç "Noghay" 333-344
Hendrik Boeschoten "Uzbek" 357-379
Claus Schönig "South Siberian Turkic" 403-417
Árpád Berta "West Kipchak Languages" 301-318
Lars Johanson "The Structure of Turkic" 30-67
- Plank, Frans "Encoding grammatical relations: acceptable and unacceptable non-distinctness" in *Historical Morphology*. ed. Jacek Fisiak. The Hague: Mouton, 1980.
- Krueger, John R. *Yakut Manual*. Bloomington, IN: Indiana University Publications, 1962.
- Micheau, Cheri. "The *voseo* in Latin America: Insights from historical sociolinguistics." *Sociolinguistics of the Spanish-Speaking World: Iberia, Latin America, United States*. ed. Carol A. Klee. Tempe, AZ: Bilingual Press, 1991.
- Hebert, Raymond J. and Nicholas Poppe. *Kirghiz Manual*. Bloomington, IN: Indiana University Publications, 1963.
- Menges, Karl H. *Qaraqalpaq Grammar*. Morningside Heights, NY: King's Crown Press, 1947.
- Mayerthaler, Willi. *Morphological Naturalness*. Tr. Janice Seidler. Ann Arbor, MI: Karoma Publishers, 1988.
- Carstairs, Andrew. *Allopmorphy in Inflexion*. London: Croom Helm, 1987.
- Trask, Larry. "A Linguistic Sketch of Basque" from
<http://www.cogs.susx.ac.uk/users/larryt/basque.sketch.html>. Dated August 23, 1996
- Wurzel, Wolfgang Ullrich. *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer, 1988.
- Carstairs-McCarthy, Andrew. *Current Morphology*. London: Routledge, 1992.
- Nadzhip, E. N. *Modern Uigur*. tr. D. M. Segal. Moscow: Nauka, 1971
- Krueger, John R. *Yakut Manual*. Bloomington, IN: Indiana University Publications, 1962.