# Linguistic Economy And The Typology of Turkic Pronouns

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

Grace Mrowicki
November 21, 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

1

connection is also realized phonologically. These concepts are generically contrasted with that of suppletion. Suppletive forms, while semantically related, lack morpho-phonological similarities. The classic example of this is the pair of *go* and *went*. Opaque forms may have phonological similarities, but they are hard to ascertain.

The idealization of the transparent, one-to-one type of system goes back to a useful, if quite vague universal principle originally set out by Humboldt, and hence referred to as Humboldt's Universal, well explained by Peirce (1955: 320) as 'the one primary and fundamental law of mental action…a tendency towards generalization." Several linguists, namely Vennemann and Hudson, understand this "tendency towards generalization" to mean that languages systematically tend towards this type of one-to-one system of specifically inflectional morphology.

An agglutinating-type morphology comes close to this type of one-to-one morphological relationship, in fact, it is the presence of exactly this kind of, once again, specifically inflectional morphology that characterizes a language as having agglutinating morphology. By way of example, consider the following group of words from Kazak, a Turkic language I will discuss in greater detail later on:

1) *zhaz* – write (root)
*zhaz-u* – writing (n., as in a piece of writing)
*zhaz-u-shy* - writer
*zhaz-u-shy-lar* - writers
*zhaz-u-shy-lar-ymyz* – our writers
*zhaz-u-shy-lar-ymyz-da* – near our writers
*zhaz-u-shy-lar-ymyz-da-ghy* – that which is near our writers
*zhaz-u-shy-lar-ymyz-da-ghy-lar-dyŋ* – of those that are near our writers

Turkic languages, a branch of the Altaic family spoken primarily in western, central and northern Asia, are characterized by their use of agglutinating morphology. This paper will primarily explore the way Turkic pronouns, in the framework of their agglutinating

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 14$^{th}$ 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 2$^{nd}$ person. In fact, it is these concepts of formality and number, especially in the 2$^{nd}$ 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 3$^{rd}$ person feminine at this time were *scho, heo, hu, ha, he*, in different dialect regions of England

| 3) FRENCH | singular | plural |
|---|---|---|
| 1st | je | nous/on[2] |
| 2nd familiar | tu | vous |
| 2nd formal | vous | |
| 3rd m./f. | il/elle | ils/elles |

| 4) RUSSIAN | singular | plural |
|---|---|---|
| 1st | ya | my |
| 2nd familiar | ty | vy |
| 2nd formal | vy | |
| 3rd m./f./n. | on/ona/ono | oni |

As examples (1), (3), and (4) show, there is a trend towards using the 2nd person plural

form to show 2nd person singular formal form.[3] This is the origin of referring to a

familiar/formal distinction as a T/V distinction (from IE examples like the French

*tu/vous*, where the familiar begins with /t/ and the formal begins with /v/), though the

term has expanded to cover more than just this method of distinguishing familiar from

formal.

There are two main ways that these IE languages make this formality distinction

in the 2nd person. The first is the 2nd plural→ 2nd singular formal, shown in (1), (3), and

(4). The second uses a 3rd person form to indicate 2nd person formal. Two familiar

examples of this are from Spanish and German:

---

[2] Because I can't ignore *on*: On is a sort of informal 1st person plural that behaves grammatically like a 3rd person singular.

[3] Brown and Gilman (1960: 254) suggest the following possible origins for the use of the 2nd pl to indicate formality, at least in Latin:

    a) First, a plural form may have been chosen for a term of respect because there were two emperors at that time. The office was unified administratively, and by using the plural pronoun in addressing one of the emperors, both could be included by implication;

    b) a second theory is that *vos* may have been the natural response to statements by the Roman emperor in which he referred to himself as *nos* (we), the sum of all the subjects of his empire.

| 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 2$^{nd}$ person formal *usted* is a reduced version of *vuestra merced* (your mercy), and behaves grammatically 3$^{rd}$ person singular pronoun. In the German example, they take advantage once again of their creative use of capitalization, and *Sie* behaves grammatically like the 3$^{rd}$ 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 2$^{nd}$ sg formal means that in the 2$^{nd}$ sg, there is a clear distinction between familiar and formal. However, it means that you can't always grammatically distinguish 2$^{nd}$ pl from 2$^{nd}$ 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 2$^{nd}$ sg. fam. *thou*, in combination with the use of the 2$^{nd}$ pl *you* to indicate 2$^{nd}$ formal, means that in Modern Standard English, *you* covers the entirety of the 2$^{nd}$ person, and there are no distinctions for number or formality anymore.

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

---

[4] Distinguished from 3$^{rd}$ sg fem *sie* because of dative form: 3$^{rd}$ pl/2$^{nd}$ form dat *ihnen*; 3$^{rd}$ sg fem dat *ihr*.

group (e.g. you lot, you guys). However, as previously mentioned, because of the very high frequency with which pronouns are used, they tend towards being phonologically compact for easier use. As such, various dialects of English have generated handier forms employing clitics or suffixes. For example:

7) Southern U.S. *y'all* (from 'you all') [jaːl]
   NJ/Philadelphia: *yous(e)* (general plural suffix /s/) [juːz] or [jʌz]
   Pittsburg/Ohio Valley: *yuns* (from you ones) [jʌnz]

*Y'all* is recognized enough in American English that it is even accepted by Word's spellchecker, which also notes that the grammatically correct conjugation of *be* for *y'all* is *are*. This indicates that it's beginning to be codified into a standard, although exact usage varies widely by region. Unfortunately, though, there's not enough room here to go into a full-fledged exploration of the evolution of distinct 2nd person plural pronouns in modern English.

To now return to Turkic languages, their pronominal paradigms also, on first glance, seem to follow this pattern of agglutination, unlike the more suppletive, or at least opaque, Indo-European examples.

Consider, for example, Turkish:

| 9) TURKISH PRONOUNS | Singular | Plural |
|---|---|---|
| 1st | ben | biz |
| 2nd familiar | sen | siz |
| 2nd formal | siz | |
| 3rd | on | onlar |

Within the context of the 1st and 2nd person[5] here, just looking at minimal differences, the pronouns appear to break down into meaningful components quite cleanly. This system

---

[5] The 3rd person pronouns in Altaic languages are derived from demonstrative pronouns (like *that*), and as such behave a little differently, but I leave them in because they are useful for comparison.

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 |
|---|---|---|
| 1st | men | biz |
| 2nd familiar | sen | sender |
| 2nd formal | siz | sizder |
| 3rd | 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/continuancy along the same place of articulation, and 1st person morphology shows a lot of allophony among the labial consonants.

in the Turkish examples. The 2<sup>nd</sup> pl familiar doesn't even have /iz/ in it. Also, the 2<sup>nd</sup> formal here does not simply use the 2<sup>nd</sup> pl form. Looking at minimal differences between the forms, we can draw the following conclusions about the 1<sup>st</sup> and 2<sup>nd</sup> person forms.

    (a) /+labial/ corresponds to 1st person
    (b) /s/ corresponds to 2<sup>nd</sup> person

However, these are the only conclusions that can be drawn about all of these forms, and only account for one of the three pieces of information that potentially need to be accounted for (person, number, and formality). Already we have a disruption in the consistency of the way that the necessary bits of meaning are spelled out that the agglutinative morphology would seem to predict, and that appears to be present in the example of Turkish. In the context of the 1<sup>st</sup> and 2<sup>nd</sup> person paradigms individually, we can draw the following additional conclusions based solely on minimal differences:

1<sup>st</sup> person:
    (a) /en/ corresponds to singular
    (b) /iz/ corresponds to plural

2<sup>nd</sup> person:
    (a) /en/ corresponds to familiar
    (b) /iz/ corresponds to formal
    (c) /der/ corresponds to plural

So it would seem that not only is there an inconsistency in the way one piece of meaning (here, plural number) is expressed, but that there is also an inconsistency in the meanings of the morphemes /en/ and /iz/, which mean different things in the context of Kazak's 1<sup>st</sup> and 2<sup>nd</sup> persons, respectively. Both of these inconsistencies work against the agglutinating morphology which is usually consistent on both of these counts. The rest of the Turkic languages, as far as the data I've seen, seem to follow either the general

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 2$^{nd}$ person pl).[7]

Because of these factors and growing inconsistencies, Turkic pronouns provide and interesting perspective on the phenomenon of 2$^{nd}$ 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 1$^{st}$ and 2$^{nd}$ 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

# Chapter 2: How far it all breaks down.

I mentioned earlier that pronouns, both conceptually and functionally, are related to inflectional morphology, because of the type of information involved (minimally person, and most times number – the same type of information involved in clearly inflectional agreement morphology). This is especially true in the case of Turkic languages, where the morphemes making up pronouns are often closely related phonologically (if not phonologically identical) to those realizing the same properties in agreement morphology. Turkic pronominal agreement morphology can be broken down into two basic types: the first is a cliticized form of the full pronoun, and so-called 'true affixes.' (see Appendix 2 for examples of each of these). The cliticized forms, unsurprisingly, have the same morphological components as well as all but identical phonological components as free pronouns. The affixal forms are also very similar structurally, although the roots which seem to correspond to person are slightly different.

Table 2.1

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

Table 2.2

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

Specfically in terms of /iz/ and /LAr/, note that they appear consistently with their presence in full pronouns of the 2nd person, which appears to demonstrate an awareness

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

("allomorphy" one notion is expressed in different words by different morphs)

III.      Many properties to one exponent syntagmatically
("overlap" one morph expresses several notions in one word)

IV.      Many properties to one exponent paradigmatically
("homonymy" one morph is used to express distinct notions in different situations)

All of these deviations can also describe the morphological situation of pronouns. In terms of the pronouns we've already discussed:

I) Spanish: Plurals in the 1st and 2nd familiar persons can be parsed out to different roots than their singular counterparts *and* include the language's general plural suffix, /s/.

II) Kazak: plural indicated in 1st person by /iz/ and in 2nd and 3rd person by /ler/

III) German (and, in fact, most of the Indo-European examples presented earlier): There is no morphological consistency between any of the nominative forms, or really anything parsable about them at all for that matter. As such, the properties that the pronoun describes can only be attributed to the pronoun as a whole, and not to any morph parsed out of the pronoun, since the word cannot be broken down any more.

IV) Standard English 2nd person. No distinction, morphological, phonological or otherwise. Not even for case.

As previously discussed, most Turkish-type paradigms do not deviate from the consistent one-to-one relationship. Kazak-type paradigms require more analysis of their transparency to be able to say exactly which deviations apply to the patterns. If nothing else, they display Deviation II in the differential expressions of plurality between the 1st person and 2nd person.

Carstairs goes on to set up a theory of memorability for inflectional morphology based on the number of pieces of information that must be memorized in order to use a given inflectional paradigm. Assuming an ideal one-to-one type of paradigm, I'll use the

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

hypothesis.' Integral in these bracketing criteria, and in Mayerthaler's treatment of transparency is the notion of 'equipotent' Brame (1974: 56). He defines two strings in a given phonological representation to be equipotent if they are identical and at least one of the two is not represented as a proper substring in the phonetic representation.

Brame's Natural bracketing Hypothesis:

(1) *Strong or production-oriented bracketing of formatives*:
A substring (a) of a string (b) can only be marked as a formative if (a) is equipotent to a string (c) and the meaning (or denotation) of (b) is a compositional function of the meaning (or denotation) of (c) and (b) – (a) (b minus a).

(2) *Weak or analysis-oriented bracketing of formatives:*
A substring (a) of a string (b) can be bracketed as a formative if (a) is equipotent to a string (c). The segment (b)-(a) corresponds to a blocked/unique formative.

The differentiation between "analysis- vs. production- oriented" formative bracketing represents one of the reflections of 'speaker morphology (is not equal to) hearer morphology'

With that in mind, Mayerthaler constructs his treatment of transparency as follows:

(a) a word is 'constructionally transparent' if individual formatives can be clearly identified according to bracketing convention (1)
e.g. peerlessness = peer + less + ness

(b) a word is 'constructionally semi-transparent' when it can be bracketed according to (2)
e.g. cranberry while it's possible to parse this as cran+berry, only *berry* has real semantic value outside of the word.

(c) a word is 'constructionally non-transparent' when neither (1) nor (2) is applicable, but word seems to contain some sort of seam in its construction
e.g. cupboard ([c^bb(e)rd]) while the written form maintains its origin as cup+board, the phonological expression of it doesn't really, such that you need to really think about it or have it pointed out to you

(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 |
|---|---|---|
| 1st | b-en | b-iz |
| 2nd | s-en | s-iz |
| 3rd | 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 |
|---|---|---|
| 1st | men | biz |
| 2nd familiar | sen | sender |
| 2nd formal | siz | sizder |
| 3rd | 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

2<sup>nd</sup> person generalizations:

      (d) /en/ corresponds to familiar

      (e) /iz/ corresponds to formal

      (f) /der/ corresponds to plural

The analysis of transparency here either has to commit to analyzing each of the three persons independently, which seems to be an undesirable approach, because, among other things, it creates formality as an inflectional category, since the 2<sup>nd</sup> person /iz/ and [der] appear in all of the inflectional agreement morphology. Assuming this analysis, it also fixes multiple deviations from the ideal one-to-one model onto the paradigm, specifically Carstairs' Deviations II, because of the distinct pluralizers for 1<sup>st</sup> and 2<sup>nd</sup> persons and IV, because of the 1<sup>st</sup> person pluralizer and the 2<sup>nd</sup> person formalizer being homonymous. This analysis begins to look even less desirable in light of Carstairs' theory of memorability, which breaks it down as follows (to make the morphemes more visible, I've replaced them with the following representations: 1<sup>st</sup> = a, 2<sup>nd</sup>= b, 3<sup>rd</sup> = c; /en/=1, /iz/=2, /ler/=3):

Table 2.7

| Kazak model 1 | sg | pl |
|---|---|---|
| 1<sup>st</sup> | a 1 | a 2 |
| 2<sup>nd</sup> familiar | b 1 | b 1 3 |
| 2<sup>nd</sup> formal | b 2 | b 2 3 |
| 3<sup>rd</sup> | c | c 3 |

This system requires a speaker to remember 6 morphemes in order to produce these 8 forms. This doesn't seem too unreasonable, except that there are a minimum of two rules that also must be memorized (unambiguously one about /en/, one about /iz/, and, debatably, one about /ler/), which brings the total up to 8, debatably 9 pieces of information necessary to produce the 8 member paradigm.

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 |
|---|---|---|
| 1st | a | e |
| 2nd familiar | b | b 1 |
| 2nd formal | c | c 1 |
| 3rd | 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

indicating plural, here it seems to have transcended plurality in a way, being borrowed lock, stock, and barrel from its origin of simply being 2nd plural into the domain of 2nd singular formal, to the extent that to recreate the distinction between singular and plural, the plural suffix must again be applied into what looks to an outside observer like a double-marking.

Table 2.9

| TUVAN PRONOUNS | singular | plural |
|---|---|---|
| 1 | men | bis |
| 2 familiar | sen | siler |
| 2 formal | siler | silerler |
| 3 | ol | olar |

The complication of the *siler* form is that it is a product of phonological change. It is non-obviously derived from *sen* + /LAr/, if Hahn's (1998: 390) claim about a similar form in Uighur can be extended to other *siler*s (Tuvan, Kyrgyz, and Uighur have them). The fact that *sen* is indiscernible in the form makes it easier to perceive *siler* as a single unit.

Table 2.10

| Tuvan model | singular | plural |
|---|---|---|
| 1 | a | e |
| 2 familiar | b | c (b 1) |
| 2 formal | c | c 1 |
| 3 | d | d 1 |

# 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),

although it is possible that other Turkic languages use this formation and the authors of the grammars I've seen haven't included descriptions of this use for whatever reason. In about half of these examples, the use of /LAr/ is described as an "individualizing plural." While I haven't seen any real description about what this really means in regards to a 1st person situation, Lars Johanson describes the use of /LAr/ generally in Turkic languages as follows: "Plural suffixes mostly signal individual plurality, e.g. Turkish *elmalar* '[single] apples'" (Johanson 1998: 51). Here it will be useful to explain a little bit how use of plural forms in Turkic languages differs from Indo-European use of it. Nouns in Turkic languages do not take a plural suffix in situations where they appear with a quantifying phrase, in situations of indefinite number, or in generic readings, as seen in the following Kazak sentences.

    (a) Men eki qorma zhejdim
        I   two persimmon eat(past)(1st sg)
     *'I ate two persimmons.'*

    (b) Biz alma zhinadyq.
        We  apple collect(past)(1st pl)
     *'We picked apples'*

    (c)   Ol uj zhasajdy
    (3rd sg) house build(gen. pres)(3rd)
    *'He/She builds houses.'*

With this understanding of the use of /LAr/, the usefulness of the "individualizing plural" becomes a little more clear. In regards to the interpretation of an individualized 1st pl pronoun, it would best be expressed as something like "each of us." What is essential to note here, though, is that this is a secondary formation, where /LAr/ is being attached to an already plural pronoun for a specific use.

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)

pl and 2$^{nd}$ formal singular (think /iz/), and the presence of the Basque general plural

suffix, /ek/ on especially the 2$^{nd}$ person pl but also on the 3$^{rd}$ person plural, and the lack of

it on the 1$^{st}$ person plural. While the vocalism of *ni*/*hi* and *gu*/*zu* may be completely

coincidental, the pattern of /ek/ use indicates a 2$^{nd}$ plural form was borrowed for use as a

2$^{nd}$ formal, and then the distinction of plural from formal was reestablished using the

normal plural morphology, which is in fact what happened (Trask, p.c.). Meanwhile, the

1$^{st}$ person has also maintained its independence of /ek/. To abstract the paradigm:

Table 3.2

| Basque model | singular | plural |
|---|---|---|
| 1 | a | e |
| 2 familiar | b | |
| 2 formal | c | c 1 |
| 3 | d | d 1 |

This is an incredibly similar pattern to that in Table 2.5, the only difference being that

there is only one 2$^{nd}$ pl form. Of the other Turkic languages not extensively discussed,

Uighur displays this abstracted pattern observed in Table 2.8.

To come back to the point I was making earlier about the 1$^{st}$ pl forms not adopting

the language's general plural suffix in the way the 2$^{nd}$ and 3$^{rd}$ persons do, I said that it

was inconsistent with Carstairs analysis of memorability in regards to number of things to

remember in order to reproduce a given paradigm. However, the fact that the Basque 1$^{st}$

pl seems relatively stable, and displays this same pattern suggests that the redundancy

that would be *guek* might not be preferable, as Carstairs' analysis would predict.

This suggests that there is some reason why the less consistent, but more compact

1$^{st}$ pl form remains as it does. Redundancy, it would appear, does count for something.

To return to my own sarcastically simplistic definition of economy, a language wants the

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 $2^{nd}$ sg formal, in addition to $2^{nd}$ 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 $1^{st}$ 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.

# Chapter 4: The more things change, the more they stay the same

The above conclusions could be interpreted as evidence for Turkic languages moving towards having a "fusional" morphology similar to that of Indo-European languages. This would play well into a classical conception of a typological cycle, wherein languages pass from isolating to agglutinating to fusional[9] (and, in some people's assessment, from here back to isolating). If nothing else, it seems an interesting case study in the kinds of interference that lead away from an idealized morphological pattern.

Typopgically, however, what's most interesting to notice is that inasmuch as the pronouns appear less transparent, the associated agreement morphology is keeping up with the pronouns. Conceptualizing this as a transition towards a fusional morphology might predict a degredation of the structural and phonological affinity between pronouns and pronominal agreement morphology. But that does not seem to be the case.

The comment it makes, however, focuses more on the morphological distinction of meaning as opposed to the compactness of the form. Simply put: the competition for most economical form by the two. On one hand, there's phonological compactness. On the other hand, there's less ambiguity. However, phonological variation being held up as the real enemy of the persistence of an idealized one-to-one morphology, some forms here, most notably the Tuvan 2nd sg formal *siler* seem to challenge the historical uber-consistency of agglutination. The form *silerler* though, indicates a commitment to clarity.

---

[9] The linearness of this especially is a little uncomfortable, since it emerged out of a time where people were using Darwinian concepts of evolution to justify all kinds of Euro-centric theory.

# 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 | -sIŋ | -sIŋdAr |
| 2 formal | -sIz | -sIzdAr |
| 3 | -dI | -dI |

| Uzbek Cliticized Pro. Agr. | Singular | Plural |
|---|---|---|
| 1 | -män | -miz |
| 2 | -sän | -siz |
| 3 | null | -lär |

| 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 |
|---|---|---|
| 1st | -m | -miz |
| 2nd | -ŋ | - ŋiz |
| 3rd | -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.