Lund University, Dept. of Linguistics  

What else Do you Need to Know about the Referents?

Karina Vamling

Abstract

In a discourse fragment, the referents are referred to by various noun phrases, pronouns and personal markers in the verbform. It is interesting to consider the substitutional character of the personal pronouns. What information do they express about the referents, i.e. what does one need to know about a referent in order to pick the right pronoun?

In a case like I know Johan. He works at our department, the pronoun he is chosen for the second reference to the person called Johan, as it corresponds to one male person, who is neither the speaker nor the hearer. The features involved are thus, person, number and semantic gender.

The selection of pronouns is somewhat different in, for instance, the Caucasian languages, where additional features of the discourse situation become relevant, such as the distance and relative location of the referents above or below the speaker/hearer.

The implications such systems may have for a formal, cross-linguistic model are briefly discussed in the paper.

BACKGROUND

Before going into detail in the system of personal and demonstrative pronouns in the Caucasian languages, some general points about systems of personal pronouns will be noted.

An early proposal about pronoun systems was made by Greenberg 1963, stating that “All languages have pronominal categories involving at least three persons and two numbers” (p. 96). A following study by Ingram 1978 proposes refinement of Greenberg’s universal: “every language designates at least four persons – I, thou, he and we” (p. 227). Ingram also states that the minimal distinction that is made in number systems is between ‘one’ and ‘more than one’. Ingram bases his generalizations on a sample of approximately 60 languages.

A larger study the typology of personal pronoun systems has been conducted by Sokolovskaja 1980, who takes as her basis personal pronouns in 400 languages. She formulates the generalization: “Each of the seven metapersons1 is expressed in the system of personal pronouns in any language” (p. 90). She

1The metapersons are defined as: $s$=I, $tu$=thou, $h$=he: 1. (I) I, 2. (you)$X$ thou, you, 3. (he)$X$ he, they, 4. (I you)$X$ we incl., 5. (I he)$X$ we excl., 6. (you he)$X$ you, 7. (I you he)$X$ we [K.V.].
KARINA VAMLING points out that when a language lacks a special form to express any of the meta-persons, its meaning is expressed together with one of the other meta-persons. For the following discussion, her generalization 16 is interesting: "In all languages a special form exists to express the 3rd meta-person." (p. 92). However, she mentions that some languages lack special personal pronouns expressing the third person. Instead, demonstrative pronouns are used. She considers such pronouns to be included in the main subsystem of personal pronouns, as they, besides their demonstrative meaning, always express the 3rd meta-person.

In a later article by Greenberg 1985 the same point is made: "In many languages there is no separate third person pronoun as distinct from one or more of the demonstratives" (p. 271). Greenberg further discusses a diachronic link found in various languages between demonstrative and personal pronouns in the third person. He observes that it is usually the distance demonstrative that develops into a personal pronoun or, if not, a demonstrative pronoun that has included the distance meaning into its range of meanings (p. 279).

Lyons 1971 also discusses the problem of demonstrative and personal pronouns found in some languages: "In many languages no distinction can be drawn between the 'demonstratives' and the 'third person pronouns'... In classical Latin (as in Greek), there was no 'third person pronoun' at all: where 'pronominal' reference was made to some 'subject of discourse' (other than the speaker or hearer), the appropriate 'demonstrative' was used – hic, iste or ille" (p. 279). The Latin demonstratives indicate (1) proximity to the speaker: hic. (2) proximity to the hearer: iste. (3) remoteness from both the speaker and the hearer: ille.

It is thus clear that a unified cross-linguistic treatment of personal pronouns in the third person poses certain problems. It does not suffice to rely on the features person and number as well as language specific features of gender.

PERSONAL PRONOUNS IN THE CAUCASIAN LANGUAGES

A striking feature of the Caucasian languages is the sharply marked difference between the first and second person as opposed to the third, i.e. the participants in the communication and the ones remaining outside. In a majority of the languages personal pronouns exist only for the first and second person whereas third person demonstrative pronouns are used as equivalents of third person personal pronouns.

The Caucasian languages

The Caucasian languages comprise some 35 languages, divided into three language families: the South Caucasian (or Kartvelian), the North-West Caucasian and the North-East Caucasian (with a major division into Nakh and Dagestan languages). It is customary to treat the Caucasian languages as a group of languages, though it should be pointed out that the genetic link between the three language families has not been satisfactorily proven (Klimov 1983). For reference, a table over the major Caucasian languages is given in Appendix 1.

Below, I will focus on some unusual and linguistically interesting features of the systems of personal/demonstrative pronouns found in the Caucasian languages. It is by no means an attempt to give an exhaustive classification.

The account below of personal and demonstrative pronouns in the Caucasian languages is based on available grammatical descriptions, not on informant work (indications of the main sources consulted for each language have been included in Appendix 1). The grammars vary considerably in detail but provide sufficient material in order to get a grasp of the relevant features on the whole.

First and second person

Proper personal pronouns are found only in the first and second persons. A large number of Caucasian languages distinguish between inclusive (including the hearer) and exclusive (excluding the hearer) forms in the first person plural. The distinction inclusive-exclusive is found in the East Caucasian languages: all the Nakh languages and among the Dagestan languages in Avar, the Andi languages and most of the Lezgian languages. It has also been suggested for Abaza among the West Caucasian languages (Genko 1955) but this claim has been disputed recently by Lomtatidze 1987:114. She argues that the form that Genko interprets as the exclusive pronoun rather is a special demonstrative form found in Abaza for first and second personal pronouns.

Another feature of personal pronouns in Abaza and the related Abkhaz is that semantic gender is distinguished in second person singular. This is also the case in personal markers on the verb, illustrated here in Abkhaz (Aristava 1982:54):

(1) a wara sara u-sa-c'az'oit
you (m.) I 2sg(m.)-lsg-talk to
‘You (male) are talking to me.’
Third person

In the third person, demonstrative pronouns are used as corresponding to personal pronouns. As seen below, the primary point of orientation in the system of demonstrative pronouns is the location of the speaker. A more unusual feature (but common in this language family) is that the location above or below the speaker is encoded in the demonstrative pronouns. The systems of personal pronouns can thus be described by a horizontal and, in some cases in addition to that, by a vertical parameter. Looking at the horizontal parameter first, one finds languages that distinguish two or three degrees of remoteness. As stated by Klimov et al. 1978:19, systems with three degrees of remoteness are found in all three language families: (a) close to the speaker, (b) close to the hearer and, (c) close to a third person. Unfortunately, the descriptions of the various languages do not use the same categories in describing these degrees, which, of course, makes a comparison between the languages difficult. Some examples from the three language families:

(2) a Lak (NE) va 
  close to the speaker
  mu 
  close to the hearer
  ta 
  far from the speaker and hearer
  (Murkelinskij JaN:497)

b Adyge (WC) mə
  close to the speaker
  mo
  visible, further from the speaker
  a
  far from both the speaker and the hearer
  (Rogava and Keraševa 1966:85)

c Georgian (SC) es
  close to first person
  eg
  close to second person
  is
  close to third pers. (Sanidze 1980:101)

Systems with two members on the horizontal parameter are less frequent. Such systems are found in all three language families, as illustrated below:

(3) a Ubykh (NW) jana close
  wana remote
  (Kumaxov JaN:698)

b Svan (SC) ala this
  eʒa that
  (Topuria JaN:83)

c Khvarsh (NE) idu least remote
  ju most remote
  (Bokarev JaN:427)

The most elaborate account of personal pronouns is given by Kibrik 1977:123-127 in his three volume description of the Dagestan language Archi. Archi belongs to the group of languages that specify location above or below the speaker, as shown in the table below:

<table>
<thead>
<tr>
<th>Noun classes:</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close to the speaker</td>
<td>jow</td>
<td>jar</td>
<td>jab</td>
<td>jat</td>
<td>jeb</td>
</tr>
<tr>
<td>Close to the hearer</td>
<td>jámu</td>
<td>jámur</td>
<td>jámub</td>
<td>jámut</td>
<td>jémim</td>
</tr>
<tr>
<td>Distant from the speaker</td>
<td>tow</td>
<td>tor</td>
<td>tob</td>
<td>tot</td>
<td>teb</td>
</tr>
<tr>
<td>Distant and above the sp.</td>
<td>Rodú</td>
<td>Rodór</td>
<td>Rodób</td>
<td>Rodót</td>
<td>Ridib</td>
</tr>
<tr>
<td>Distant and below the sp.</td>
<td>gudú</td>
<td>godór</td>
<td>godób</td>
<td>godót</td>
<td>gidib</td>
</tr>
</tbody>
</table>

Some examples where these pronouns are used as demonstratives and personal pronouns (Kibrik 1977:169, 186; 165, 171):

(4) a jat noŋ'axut těnši uqla
  'Walk through this (close to the speaker) room'
Archí, like most of the East Caucasian languages, distinguishes a number of noun classes, as is shown by the five columns in Table 1. The Archí system differentiates four classes in the singular and has one common class in the plural. The noun class markers do not show up on the nouns themselves but on adjectives (dōzu-v, dōzu-r, dōzu-b, dōzu-t ‘big’), pronouns (cf. Table 1), numerals (qve-vu, qve-ru, qve-bu, qve-t’t ‘two’) and verbs (so-v-k:as, sa-r-k:as, sa-b-k:as, sa-o-k:as ‘look’).

Table 1. Noun classes in Archí

<table>
<thead>
<tr>
<th>Members in the class</th>
<th>markers (in the sing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Nouns denoting men</td>
</tr>
<tr>
<td>II</td>
<td>Nouns denoting women</td>
</tr>
<tr>
<td>III</td>
<td>Animals, crops etc.</td>
</tr>
<tr>
<td>IV</td>
<td>Young/small animals, metals etc</td>
</tr>
</tbody>
</table>

Table 2. Noun classes in Archí (Xajdakov Jarn 1967:611)

According to Kibrik 1977:126, the same range of semantic distinctions that is found in demonstrative pronouns is also present when these pronouns are used as personal pronouns. When the pronouns are used anaphorically, the class marker is chosen on the basis of the class of the antecedent. When it is used deictically, the class of the pronoun is chosen corresponding to the person or thing referred to. If the referent is unknown, the fourth class is used.

The vertical parameter specifying location above or below the speaker is, apart from Archí, found in Lezgi, Tabasaran, Agul, Avar, Andi, Karata, Akhvakh, Tindi, Chamalal, Lak and Dargva. One would have to assume that the life of the peoples of Dagestan in a severe mountainous region has contributed to this feature. The Archí example represents a system where the vertical parameter is limited to the pronouns denoting maximal distance from the speaker. In some other languages, such as Akhvakh and Tindi, this applies also to a degree closer to the speaker. As seen in Table 3, Akhvakh has two forms for ‘this’ and ‘that’, that combine with markers corresponding to the meanings ‘above’, ‘at same level’ and ‘below’: l-(nI), d-, g- (Magomedbekova Jarn:342).

Table 3. Third person (demonstrative) pronouns in Akhvakh

<table>
<thead>
<tr>
<th></th>
<th>‘This’</th>
<th>‘That’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above</td>
<td>ha-ve</td>
<td>hu-ve</td>
</tr>
<tr>
<td>Same level</td>
<td>ha-la-ve</td>
<td>hu-lu-ve</td>
</tr>
<tr>
<td>Below</td>
<td>ha-da-ve</td>
<td>hu-du-ve</td>
</tr>
<tr>
<td></td>
<td>ha-ge-ve</td>
<td>hu-gu-ve</td>
</tr>
</tbody>
</table>

Two sets of terms for the vertical parameter appear to be the maximum. The Rikvalin dialect of Andi is quoted by Xajdakov 1980:38 as having the most elaborate system. It differentiates three horizontal degrees but has vertical terms only in two of them.

Table 4. Third person (demonstrative) pronouns in Andi (Rikvalin dialect), Class I

<table>
<thead>
<tr>
<th></th>
<th>Close to sp.</th>
<th>Further away</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above</td>
<td>hel’e-v</td>
<td>hunl’o-v</td>
<td></td>
</tr>
<tr>
<td>Same level</td>
<td>ho-v</td>
<td>hev</td>
<td>hundo-v</td>
</tr>
<tr>
<td>Below</td>
<td>hege-v</td>
<td>huno-v</td>
<td></td>
</tr>
</tbody>
</table>

Thus, the horizontal parameter defines two or three degrees. The two spatial parameters, the horizontal and vertical, combine at one or two points or not at all.

A recurring statement in the descriptions of the various Caucasian languages that I have examined, is that demonstrative pronouns are used in the function of third person personal pronouns. In some cases one set of demonstrative pronouns is given as corresponding to the third person personal pronouns, in other cases the author points out that there are many forms that occur as the equivalents of...
third person of personal pronouns. As predicted by Greenberg 1985:279, in the
cases where one set of demonstratives has developed into the third person of the
personal pronouns, it is the set that corresponds to the neutral, remote
demonstratives. This is the case in, for instance, Georgian.

FORMALIZING THE SYSTEMS
In an attempt to generate fragments of natural language, it appears inevitable to
include a representation of the discourse situation into which these fragments are
embedded. This point is emphasized by Sigurd 1987:150 in relation to problems
of automatic translation within the framework of Referent Grammar:

The choice of the pronouns, which depends on the features (number, gender, as e.g.
French le/là, les) poses a knotty problem, however. A dog (chien) would require a
following 'le', while a flower (fleur) would require 'la', both corresponding only to
English 'it'. The dog (hund) requires 'den' in Swedish, a child (barn) would require
'det' as a pronoun. One way to handle this situation is to use text referent grammar,
keeping track of the referents in the text and the information about them.

As the situation in the languages discussed here is even more complex, where the
personal and spatial deictic features play a most prominent role, the information
connected with the discourse referents has to be highly structured. I propose a
framework along the following lines, in order to represent the relevant
components of the discourse situation:

(6) Speaker: \[Ref\]  
Hearer: \[Ref\]  
Time: \[Tref\]  
Utterance: ‘...’

Each referent has to be identified by a unique index. The location of the referents
has to be specified as well as the time reference. Finally, semantic information
such as gender etc. has to be included in the representation.

(7) Speaker: \[Ref a\]  
\[Loc L1\]  
\[Gen fem\]  
Hearer: \[Ref b\]  
\[Loc L2\]  
\[Gen m\]

An utterance that fits into a situation like the one outlined above in (7), has a
subject in the first person singular and a direct object in the third person
singular, corresponding to a male person being close to the hearer. The verb
form is required to be in the present, as the time reference of the utterance and
the situation described by the utterance are chosen to be the same.

The kind of personal and spatial information that is accessible in a
representation like this, is needed for the selection of personal cross-referencing
affixes as well as other verbal affixes. I will illustrate this with a few examples.

Most of the North-East Caucasian languages are characterized by class
agreement. This means that a verbform reflects the class (and number) of the
subject or object. The examples below are taken from Dargva, which differenti­
ates three classes (Abdullaev 1971:71):

<table>
<thead>
<tr>
<th>Class</th>
<th>Subject</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>nu Ø-uzulra</td>
<td>I work</td>
</tr>
<tr>
<td>II</td>
<td>nu r-uzulra</td>
<td>‘-‘</td>
</tr>
<tr>
<td>III</td>
<td>nu b-uzulra</td>
<td>‘-‘</td>
</tr>
</tbody>
</table>

Second person sg  
<table>
<thead>
<tr>
<th>Class</th>
<th>Subject</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>fiu Ø-uzulri</td>
<td>You work</td>
</tr>
<tr>
<td>II</td>
<td>fiu r-uzulri</td>
<td>‘-‘</td>
</tr>
<tr>
<td>III</td>
<td>fiu b-uzulri</td>
<td>‘-‘</td>
</tr>
</tbody>
</table>

The Adyge verbal affix ‘qa represents another kind of spatio-personal
dependence. It is used to indicate that the action denoted by the verbform is
directed towards the participants in the conversation (Rogava and Keraševa
1966:113). Compare the following examples (the morphological setup of the
verb is Sub-(‘qa)-Obj-root-(tense)):

(9) a s-Ø-e-pl’a I look at him  
wa-Ø-e-pl’a You look at him  
Ø-Ø-e-pl’a He looks at him
A similar alternation is found in the two Georgian preverbs *mi* and *mo*, where *mi* marks direction from the speaker (and hearer) and *mo* the direction towards the speaker (Sanidze 1980:240).

(10) a  *micuravs – ikitk’en sadac me ar var*
    ‘he swims – from a place, where I am not’

b  *mocuravs – aketk’en, sadac me var*
    ‘he swims – towards the place where I am’

Other examples of the same alternation, but where the action is directed towards the addressee (indirect object), are for instance the Georgian verbs below (Sanidze 1980:246):

(11) a  *mo-m-it’an-a*
    mo-1sgIO-fetch-Aor
    ‘he fetched it to me’

b  *mi-u-t’an-a*
    mi-3sgIO-fetch-Aor
    ‘he fetched it to him’

The conditions on use of the preverb *mo* are met in the representation (12) below:

(12) Speaker:  [Ref a][Loc L1]

Hearer:  [Ref b][Loc L2]

Time:  [Tref t1]

Summing up, the aim of this paper has been to point at several areas – personal pronouns and verbal affixes as well as other orientational markers – where a standard treatment using only person/number categories is not sufficient. It is necessary to take into consideration orientational parameters as well as semantic information about the referents in the discourse. Data from various Caucasian languages have been used to illustrate these points. A framework has been proposed that incorporates such spatio-personal features of the discourse situation where the utterance occurs.

REFERENCES


Bokarev, Jevgenij A. 1959. *Cezskie (didojiske) jazyki Dagestan*.


Appendix I. The Caucasian Languages

South Caucasian Languages
- Georgian
- Mingrelian
- Svan

North-West Caucasian Languages
- Abkhaz
- Abaza
- Adyghe
- Kabardian
- Ubykh (now extinct)

North-East Caucasian Languages
1. Nakh Languages
   - Chechen
   - Ingush
   - Bats (or Tsova-Tush)
2. Dagestan Languages
   - Avar
   - Andi Languages
     - Andi
     - Botlikh
     - Godoberi
     - Karata
     - Akhvakhi
     - Bagval
     - Tindi
     - Chamalal
   - Tsez Languages
     - Tsez
     - Khvarsh
     - Ginukh
     - Bezhiti
     - Gunzib
   - Lak
   - Dargwa

Lezgian Languages
- Lezgi
- Tabasarani
- Agul
- Rutul
- Tsachur
- Archi
- Kryz
- Budukh
- Khinalug
- Udi

Main sources
- Şanidze 1980
- Kiziria JaN
- Topuria JaN
- Lomtatidze JaN
- Lomtatidze JaN, Genko 1955
- Rogava and Keraševa 1966
- Šagirov JaN
- Kumaxov JaN
- Dešeriev JaN
- Dolakova JaN
- Dešeriev 1953; JaN
- Madjiev JaN
- Tserovszadze JaN
- Gudava JaN
- Gudava JaN
- Megomedbevka JaN
- Megomedbevka JaN
- Gudava JaN
- Gudava JaN
- Bokarev 1949, Megomedbevka JaN
- Bokarev 1959
- Bokarev 1959
- Bokarev 1959
- Bokarev 1959
- Žirkov 1955, Murkelinskij JaN
- Abdullahov JaN
- Mejlanova JaN
- Žirkov 1948, Xanmagomedov JaN
- Megomedov JaN
- Džejranitišilži JaN
- Talibov JaN
- Kibrik 1971, Xajdakov JaN
- Saadiev JaN
- Dešeriev JaN
- Kibrik 1972, Dešeriev JaN
- Panvilidine JaN