Sherard, Michael. 1980. *A synchronic phonology of modern colloquial Shanghai* (=*Computational Analyses of Asian & African Languages* 15). Tōkyō: National Inter-University Research Institute of Asian and African Languages and Cultures.

Sigurd, Bengt. 1961. 'The code shift in Old Norse'. *Studia Linguistica* 15, 10-21.

Sigurd, Bengt. 1965. *Phonotactic structures in Swedish.* Lund: Uniskol.

Svantesson, Jan-Olof. 1983. *Kammu phonology and morphology.* Lund: Gleerup.

Svantesson, Jan-Olof. 1985. 'Vowel harmony shift in Mongolian'. *Lingua* 67, 283-327.

Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change.* Berlin: Mouton de Gruyter.

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 meta-persons[1] is expressed in the system of personal pronouns in any language" (p. 90). She

[1]The metapersons are defined as: я=I, ты=thou, он=he: 1. (I)^X 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.].

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°až°oit
         you (m.) I   2sg(m.)-1sg-talk to
         'You (male) are talking to me.'

(1) b　sara bara　s-ba-c°až°oit
　　　I　you (f.)　1sg-2sg(f.)-talk to
　　　'I am talking to you (female).'

　c　sara wara　s-wa-c°až°oit
　　　I　you (m.)　1sg-2sg(m.)-talk to
　　　'I am talking to you (male).'

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. (Šanidze 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)　　jɔna　　close
　　　　　　　　　　　　wana　　remote
　　　　　　　　　　　　　　　　(Kumaxov JaN:698)

　b　Svan (SC)　　　ala　　this
　　　　　　　　　　　eža　　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 1. Third person pronouns in Archi

Noun classes:	I	II	III	IV	Plural
Close to the speaker	jow	jar	jab	jat	jeb
Close to the hearer	jámu	jámur	jámub	jámut	jémim
Distant from the speaker	tow	tor	tob	tot	teb
Distant and above the sp.	Rodú	Rodór	Rodób	Rodót	Ridíb
Distant and below the sp.	gudú	godór	godób	godót	gidíb

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

(4)　a　*jat* noǩ'axut tēnši uqIa
　　　　'Walk through *this* (close to the speaker) room'

(4) b   ez koat'si it'u *jat*
          'I don't need *it* (close to the speaker) now.'

(5) a   *jamut* q'onq' un hinĭtiš x̄a?
          'From where did you take *this* (close to the hearer) book?'

     b   *jamur* q'ardili jatma
          '*She* (close to the hearer) was sitting on the top.'

Archi, like most of the East Caucasian languages, distinguishes a number of noun classes, as is shown by the five columns in Table 1. The Archi 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 (*dözu-v, dözu-r, dözu-b, dözu-t* 'big') , pronouns (cf. Table 1), numerals (*qve-vu, qve-ru, qve-bu, qve-t'u* 'two') and verbs (*so-v-k:as, sa-r-k:as, sa-b-k:as, sa-Ø-k:as* 'look').

Table 2. Noun classes in Archi (Xajdakov JaN 1967:611)

	Members in the class	markers (in the sing)
I	Nouns denoting men	v, u
II	Nouns denoting women	d, r, Ø
III	Animals, crops etc.	b, v, Ø
IV	Young/small animals, metals etc	t, t', Ø

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 Archi, 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 Archi 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-(лыI), d-, g-* (Magomedbekova JaN:342).

Table 3. Third person (demonstrative) pronouns in Akhvakh

	'This'	'That'
	ha-ve	hu-ve
Above	ha-la-ve	hu-lu-ve
Same level	ha-da-ve	hu-du-ve
Below	ha-ge-ve	hu-gu-ve

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

	Close to sp.	Further away	Remote
Above		hel'e-v	hunl'o-v
Same level	ho-v	hev	hundo-v
Below		hege-v	hungo-v

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/la, 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]

     Time: [Tref   t1]

     Utterance: '...'
               S → SUB,   DO,   V
               [Ref a] [Ref   c] [Tref =t1]
                        [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 differentiates three classes (Abdullaev 1971:71):

(8) First person sg   nu Ø-uzulra   I work   Class I
                        nu r-uzulra   -"-   Class II
                        nu b-uzulra   -"-   Class III
     Second person sg   fiu Ø-uzulri   You work   Class I
                        fiu r-uzulri   -"-   Class II
                        fiu b-uzulri   -"-   Class III

The Adyge verbal affix 'qə 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-('qə)-Obj-root-(tense)):

(9) a   s-Ø-e-pl'ə   I look at him
          wa-Ø-e-pl'ə   You look at him
          Ø-Ø-e-pl'ə   He looks at him

b   Ø-'qə-sa-pl'ə   He looks at me
     Ø-'qə-wa-pl'ə   He looks at you

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 (Šanidze 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 (Šanidze 1980:246):

(11) a   mo-m-it'an-a   mo-m-c'er-a
         mo-1sgIO-fetch-Aor   mo-1sgIO-write-Aor
         'he fetched it to me'   'he wrote it to me'

     b   mi-u-t'an-a   mi-s-c'er-a
         mi-3sgIO-fetch-Aor   mi-3sgIO-write-Aor
         'he fetched it to him'   'he wrote 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]

Utterance: 'mo-m-c'er-a c'erili'
              (He wrote me a letter.)
              S → SUB, DO,   IO   V
              [Ref c] [Ref   d] [Ref a] [Tref t2 before 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 neccessary 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

Abdullaev, Zapir G. 1971. *Očerki po sintaksisu darginskogo jazyka*. Moskva: Nauka. [The syntax of Dargva].

Aristava, Šota K. 1982. *Problema prostogo predloženija v abxazskom jazyke*. Tbilisi: Izd. tbilisskogo universiteta. [The simple sentence in Abkhaz].

Bokarev, A. A. 1949. *Očerk grammatiki čamalinskogo jazyka*. Moskva: Izd. AN SSSR. [Grammar of Chamalal]

Bokarev, Jevgenij A. 1959. *Cezskie (didojskie) jazyki Dagestana*. Moskva: Izd. AN SSSR. [The Tsez (Didoj) languages of Dagestan].

Dešeriev, Ju. D. 1953. *Bacbijskij jazyk. Fonetika. Morfologija. Sintaksis. Leksika*. Moskva: Izd. AN SSSR. [The Bats language. Phonetics, morphology, syntax and lexicon].

Genko, A. N. 1955. *Abazinskij jazyk. Grammatičeskij očerk narečija tapanta*. Moskva: Izd. AN SSSR. [Grammar of Abaza (Tapanta dialect)].

Greenberg, Joseph. 1963. 'Some universals of grammar with particular reference to the order of meaningful elements'. *Universals of language*, ed. Joseph Greenberg, 73-113. Cambridge, Mass.: MIT Press.

Greenberg, Joseph. 1985. 'Some iconic relationships among place, time, and discourse deixis'. *Iconicity in Syntax.*, ed. John Haiman, 271–287. Amsterdam: John Benjamins.

Ingram, David. 1978. 'Typology and Universals of Personal Pronouns', *Universals of Human Language*, ed. Joseph Greenberg, Vol. 3, 213-247. Stanford: Stanford University Press.

JaN=*Jazyki narodov SSSR. Iberijsko-kavkazskie jazyki*, Vol. 4. Moskva: Nauka, 1967. [The languages of the USSR: The Ibero-Caucasian languages].

Kibrik, Aleksandr E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka. Vol 2: Taksonomičeskaja grammatika.* Moskva: Izd. Moskovskogo universiteta. [The structure of Archi].

Kibrik, Aleksandr E., Sandro V. Kodzasov and I. P. Olovjanikova 1972. *Fragmenty grammatiki xinalugskogo jazyka.* Moskva: Izd. Moskovskogo universiteta. [Grammar of Khinalug].

Klimov, Georgij A. 1983. 'Voprosy genetičeskix, tipologičeskix i areal'nyx vzaimootnošenij kavkazskix jazykov'. *Genetičeskie, areal'nye i tipologičeskie svjazy jazykov Azii,* ed. Vadim M. Solncev, 34–41. Moskva: Nauka. [Genetic, areal and typological relations between the languages of Asia].

Klimov, Georgij A. et al. 1978. *Strukturnye obščnosti kavkazskix jazykov.* Moskva: Nauka. [Structural features shared by the Caucasian languages].

Lomtatidze, Ketevan. 1987. 'K voprosu ob inkljuzive i ekskluzive v abxazskom i abazinskom jazykax'. *Annual of Ibero-Caucasian Linguistics.* XIV. Tbilisi: Mecniereba, 113–115. [About the inclusive and exklusive distinction in Abkhaz and Abaza].

Lyons, John. 1971. *Introduction to Theoretical Linguistics.* Cambridge: Cambridge University Press.

Rogava, Georgi V. and Z. I. Keraševa 1966. *Grammatika adygejskogo jazyka.* Krasnodar: Krasnodarskoe knižnoe izd. [Grammar of Adyge].

Šanidze, Akaki. 1980. *kartuli enis gramat'ik'is sapudzvlebi.* Selected works in 12 volumes. Vol. 3. Tbilisi: Izd. tbilisis universit'et'is gamomtsemloba. [Fundamentals of the grammar of the Georgian language].

Sigurd, Bengt. 1987: 'Referent Grammar (RG) in computer comprehension, generation and translation of text (SWETRA)'. *Working Papers* 31, 137–167. Lund: Dept. of Linguistics.

Sokolovskaja, N. K. 1980. 'Nekotorye semantičeskie universalii v sisteme ličnyx mestoimenij'. *Teorija i tiplogija mestoimenij,* ed. Igor' F. Vardul', 84-103. Moskva: Nauka. [Theory and typology of pronouns].

Xajdakov, Said. M. 1980. *Principy imennoj klassifikacii v dagestanskix jazykax.* Moskva: Nauka. [Principles of noun classification in Dagestan languages].

Žirkov, L. I. 1955. *Lakskij jazyk. Fonetika i Morfologija.* Moskva: Izd. AN SSSR. [Phonetics and morphology of Lak].

Žirkov, L. I. 1948. *Tabasaranskij jazyk. Grammatika i teksty.* Moskva: Izd. AN SSSR. [Tabasaran grammar and texts].

Appendix 1. The Caucasian Languages

	Main sources
South Caucasian Languages	
Georgian	Šanidze 1980
Mingrelian	Kiziria JaN
Svan	Topuria JaN
North-West Caucasian Languages	
Abkhaz	Lomtatidze JaN
Abaza	Lomtatidze JaN, Genko 1955
Adyge	Rogava and Keraševa 1966
Kabardian	Šagirov JaN
Ubykh (now extinct)	Kumaxov JaN
North-East Caucasian Languages	
1. Nakh Languages	
Chechen	Dešeriev JaN
Ingush	Dolakova JaN
Bats (or Tsova-Tush)	Dešeriev 1953; JaN
2. Dagestan Languages	
Avar	Madieva JaN
Andi Languages	
Andi	Tsertsvadze JaN
Botlikh	Gudava JaN
Godoberi	Gudava JaN
Karata	Magomedbekova JaN
Akhvakh	Magomedbekova JaN
Bagval	Gudava JaN
Tindi	Gudava JaN
Chamalal	Bokarev 1949, Magomedbekova JaN
Tsez Languages	
Tsez	Bokarev 1959
Khvarsh	Bokarev 1959
Ginukh	Bokarev 1959
Bezhiti	Bokarev 1959
Gunzib	Bokarev 1959
Lak	Žirkov 1955, Murkelinskij JaN
Dargva	Abdullajev JaN
Lezgian Languages	
Lezgi	Mejlanova JaN
Tabasaran	Žirkov 1948, Xanmagomedov JaN
Agul	Magometov JaN
Rutul	Džejranišvili JaN
Tsachur	Talibov JaN
Archi	Kibrik 1977, Xajdakov JaN
Kryz	Saadiev JaN
Budukh	Dešeriev JaN
Khinalug	Kibrik 1972, Dešeriev JaN
Udi	Pančvidze JaN