Variability in noun classes assignment in Zilo Andi: experimental data[1]

George Moroz, Samira Verhees
Linguistic Convergence Laboratory at NRU HSE, Moscow
agricolamz@gmail.com, jh.verhees@gmail.com

Abstract

This paper evaluates the inter speaker variation in noun class assignment among speakers of the Zilo dialect of Andi (a Nakh-Daghestanian language spoken in the Republic of Daghestan). The nominal lexicon in Andi is divided in three to six classes, depending on the dialect. In dialects with more numerous classes, there are two to three classes for inanimate objects with no obvious semantic distinction between them, while the remaining three classes (male, female, non-human animate) are semantically transparent and predictably refer to either male, female or non-human animate referents respectively. We designed an experiment to test whether the assignment of inanimate noun classes is consistent across speakers in different layers of the lexicon, including native words, older loan words, and more recent borrowings from Russian. As we will show, speakers are fairly consistent in assigning certain noun classes, though some variation occurs in all layers of the lexicon; variation is considerably higher with respect to more recent loan words and among younger speakers.

1. Introduction

Andi is a minority language spoken in a mountainous area in northwestern Daghestan (one of the North Caucasian autonomous republics of the Russian Federation). The nominal lexicon in Andi is divided into three to six classes. Some of these classes are semantically transparent, while others are not; in addition to masculine and feminine, non-human animates and two to three kinds of inanimate objects can be distinguished. The difference between the various inanimate classes remains unclear. All nouns in Andi belong to a class. This includes the influx of recent borrowings from Russian connected to technological innovations. Surprisingly, newly introduced nouns for objects like computers and simcards are assigned *different* inanimate classes, which suggests some kind of semantic distinction between these classes. As this regularity seems far from straight-forward, we decided to conduct an experiment to measure inter speaker variation regarding noun class assignment. For this purpose, we compiled a list of 114 lexical items, which includes borrowings from different periods as well as basic lexicon, and verified their class affiliation with sixteen speakers of different ages and genders. We recorded variation in every part of the lexicon, though it was significantly higher for borrowed words. We also detected a stronger tendency to variation among younger speakers. In Section 2 of this paper we introduce the Andi language, its dialects and sociolinguistic background. Section 3 provides a brief introduction to the grammatical category of noun class and the class systems attested in various dialects of Andi, as well as those of neighbouring and closely related languages. In Section 4 we discuss the pure agreement percentages among the surveyed speakers and the results from Fleiss' Kappa test. Section 5 summarizes our findings.

1 This research was conducted as part of the project "The Andi morphosyntax in a typological perspective" supported by the Russian Foundation for Basic Research (RFBR), grant No. 18-012-00852a.

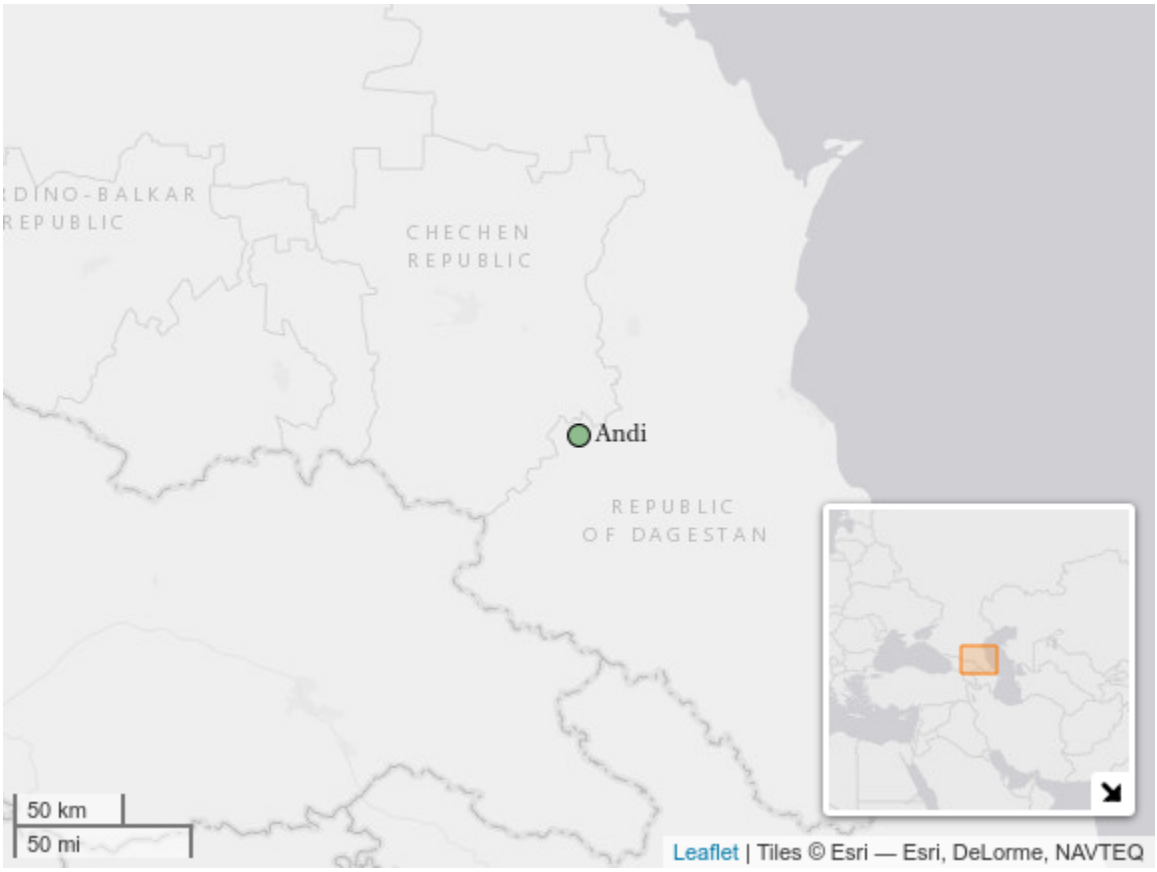

Figure 1. Andi on the map

2. Andi (background)

Andi is an unwritten language of Daghestan (see Figure 1). Together with Akhvakh, Karata, Botlikh, Godoberi, Chamalal, Bagvalal and Tindi, it makes up the Andic branch of the Nakh-Daghestanian (or East Caucasian) language family. The Andic branch is usually subsumed under a macro-branch it shares with Avar (Schulze 2013), which is also an influential neighbour of the Andic languages. Andi is spoken in the Botlikh district of the Republic of Daghestan, which borders on the Chechen republic. From some of the Andi villages Chechnya can be reached by foot. All of the Andic languages are spoken in a rather compact area, wedged in between large clusters of Chechen and Avar speaking villages (see Figure 2). Each of the nine Andi speaking villages is considered to have a distinct dialect with some minor idiosyncrasies. These dialects can be divided in two major groups: **Upper**, which includes the Andi, Gagatli, Gunkha, Rikvani, Ashali, Chanko and Zilo dialects, and **Lower**, which includes only the Muni and Kwanhidatli dialects (Tsertsvadze 1965). The names of the groups refer to their geographical location. The Upper villages are situated higher in the mountains, while the Lower dialects are spoken at the foot of these mountains, along the Andi Koisu river bank. There is mutual intelligibility among dialects, although speakers of dialects from the Upper group have trouble understanding speakers of Lower dialects (in particular that of Muni) and vice versa. Hence, they prefer to communicate in Avar, while within dialect groups they speak their native language. In this paper, we consider material from the village Zilo (see Figure 2). Although Zilo is considered one of the Upper dialects, it shares some features with the Lower group (such as the *-e* habitual). As the map

shows, Zilo is surrounded almost exclusively by Andi villages, with the exception of some Avar villages in the east. There are also Botlikh, Karata and Chechen villages in the neighbourhood.

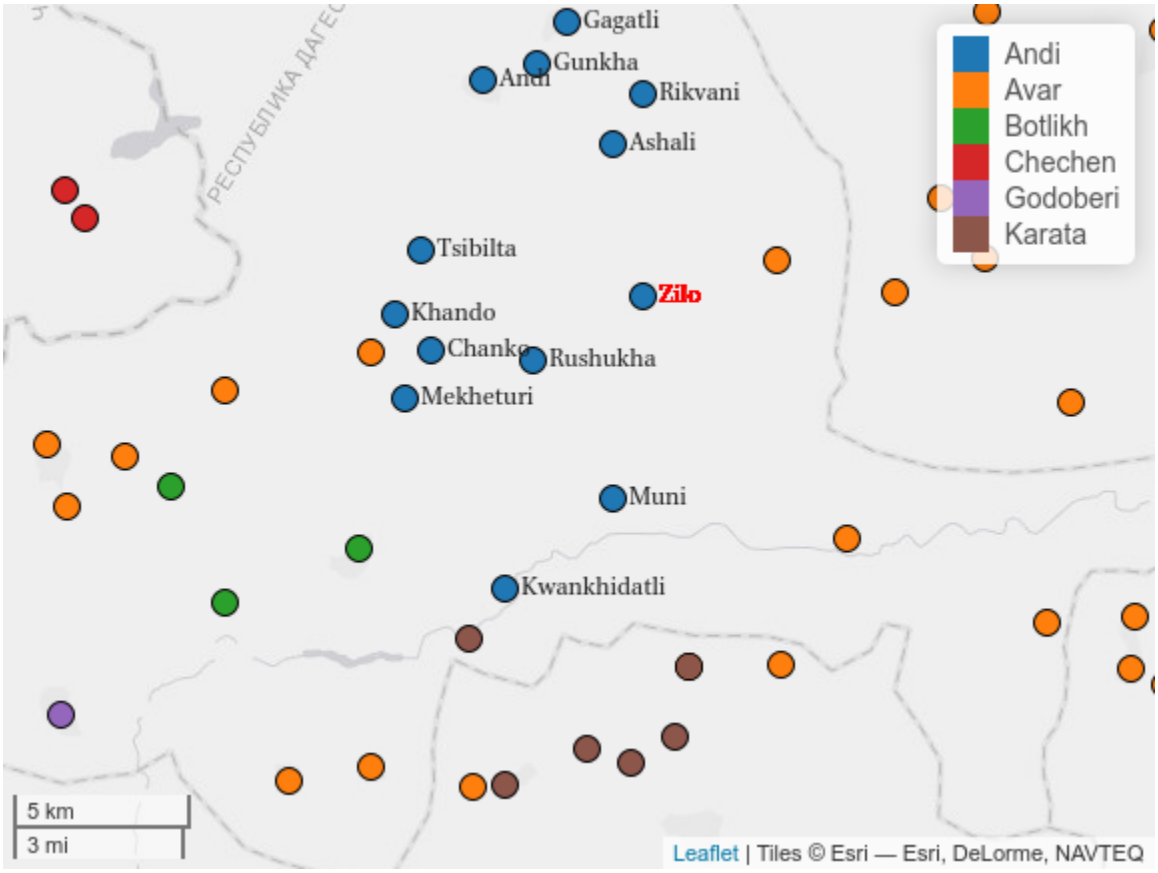

Figure 2. The Andi villages (labeled) and their neighbours

Almost all speakers of Andi and other Andic languages speak Avar and Russian in addition to their mother tongue. Avar was the major local lingua franca in this area before the introduction of Russian and, to a certain extent, still is. Speakers of Andic (as well as speakers of Tsezic) languages use it to communicate with each other and with local Avars. The language is also taught in school during the course called "native language" (*rodnoy yazyk* in Russian). For speakers of Andi, Avar is the language of literacy. Before the introduction of modern technology like computers and cellphones, they would write letters to each other in Avar. Currently, they use the extended Cyrillic script of Avar (featuring, among other things, special diacritics for ejective consonants) to write their own language. Russian was introduced as the language of administration and education in the first half of the 20th century and subsequently has strengthened its position throughout the republic, cf. (Dobrushina 2011). Andi speakers use it to communicate with people from other regions of Daghestan where Avar is not spoken.

As a result of this multilingualism, the Andi lexicon contains multiple layers of borrowings from different languages and different periods. Besides loans from Avar and Russian this includes religious terminology from Arabic, and cultural borrowings from Arabic, Persian and Kumyk (a local Turkic

language), some of which were evidently borrowed through Avar. The names of all the technical innovations of the 20th and 21st centuries were taken from Russian with little to no phonetic adaptation. In Russian these terms were mostly borrowed and adapted from English. The lack of phonetic adaptation in more recent Russian loans is connected to the increasingly important role of the Russian language in Daghestan. Earlier loans were subjected to some phonetic change, cf. *kempeti* (*kanfeti*, 'sweets'), *tʃajnik'i* (*tʃajnik* 'kettle'), and *ʃ:igolati* (*ʃ:akalat* 'chocolate' (Salimov 2010 (1968): 43), as opposed to more recent words (see Section 4). Borrowed words often have some phonological, orthographic, morphological or syntactic features in the donor language which are incompatible with the system of the recipient language. In this paper, we explore how recent Russian borrowings are integrated in the Andi system of noun classes. There is only one related study on similar experiment in other Nakh-Daghestanian language — Tsezic (see (Gagliardi and Lidz 2014)), but there are a lot of similar experiments based on other languages (e. g. (Stolz 2009)).

3. Noun class as a grammatical category

Noun class is one of the central grammatical categories of the Nakh-Daghestanian languages. It divides the nominal lexicon into varying numbers of classes. We prefer the term noun class to gender, because in Andi in particular, the division of classes is only partially semantically transparent. The masculine and feminine noun classes are assigned to nouns describing male and female (human) referents respectively, based on their biological gender. Everything else may be assigned one general non-human class, or be divided into several different classes, depending on the dialect. As we will show later, the various classes that may be assigned to objects do not show a clear semantic distinction.

A distinct class is assigned to each noun, with the exception of a small group of roots which adapt their class to the gender of the actual referent, for example:

- *ʃ:u-w ketʃ'eqan* good-m singer 'a good singer (male)'
- *ʃ:u-j ketʃ'eqan* good-f singer 'a good singer (female)'

Noun class is a covert category. Only a few nominal roots contain a class marker, e.g. *CL-ots:i* 'siblings'. Agreement is expressed on verbs, pronouns, adjectives and some other types of attributes. Demonstratives form the only part of speech which always carries a class agreement slot. For all other parts of speech the presence of an agreement slot is lexically determined: some roots have it, while others do not. On verbs the slot is usually prefixal (and, rarely, infixal). Adjectives can have a prefixal, infixal or suffixal slot, and demonstratives and pronouns always feature a suffixal slot. Noun class as an agreement class is partly merged with number in Andi (see below).

Table 1. Noun class systems in dialects of Andi

			m[2]	f	an	¬an 1	¬an 2	¬an 3	¬h
(Alekseev 1999: 221)	Andi	sg	w	j	b	b	r	-	-
(Alekseev 1999: 221)	Andi	pl	w	j	j	b	r	-	-
(Salimov 2010 (1968): 47–62)	Gagatli	sg	w	j	b	b	r	-	-

2 We use the following notation: m — masculine, f — feminine, an — animate beings (non-human), ¬an — inanimate objects (which may include some insects) and ¬h — general non-human (including both animate and inanimate).

(Salimov 2010 (1968): 47–62)	Gagatli	pl	w	j	j	b	r	-	-
(Suleymanov 1957: 131–162)	Rikvani	sg	w	j	b	b	r	b	-
(Suleymanov 1957: 131–162)	Rikvani	pl	w	j	j	b	r	r	-
Field data	Zilo	sg	w	j	b	b	r	-	-
Field data	Zilo	pl	w	j	j	b	r	-	-
(Tsertsvadze 1965)	Muni	sg	w	j	-	-	-	-	b
(Tsertsvadze 1965)	Muni	pl	w	j	-	-	-	-	b

Table 1 illustrates that the noun class system can be structurally different in different dialects. In the Muni dialect (as well as in Kwankhidatli) as described by (Tsertsvadze 1965), there are only three classes (masculine, feminine and "neuter", or "everything else"). The other dialects differentiate non-human animates (i.e. animals) and two to three classes of inanimate objects. Rikvani has the most elaborate system with a total of three inanimate classes, although the third inanimate class contains only a handful of nouns. The Andi system differs radically from that of the neighbouring languages, especially those which are more closely related; they feature smaller systems with a general non-human class, similar to the Lower Andi dialects. Andi seems to be more similar to Chechen in this respect, but the resemblance is limited to the number of classes; the markers themselves and the structure of the system are different. Because class is partly fused with number, the amount of noun classes (or the amount of classes distinguished by a specific prefix) may differ in plural. Languages like Botlikh, for example, distinguish three classes in singular (masculine, feminine and neuter), and only two in plural (human vs. non-human). In Andi it is rather the other way around: some classes have the same marker in singular (e.g. animate and inanimate 1), but we consider them distinct classes because they are differentiated in plural.

Table 2. Class inventory of languages spoken in neighbouring villages

			m	f	¬h	¬h_1	¬h_2	¬h_3
(Alekseev and Atayev 1997: 41)	Avar	sg	w	j	b	-	-	-
(Alekseev and Atayev 1997: 41)	Avar	pl	r/l	r/l	r/l	-	-	-
(Magomedbekova 1971: 41)	Karata	sg	w	j	b	-	-	-
(Magomedbekova 1971: 41)	Karata	pl	b	b	r	-	-	-
(Gudava 1962: 253–254)	Botlikh	sg	w	j	b/m	-	-	-
(Gudava 1962: 253–254)	Botlikh	pl	b	b	r/n	-	-	-
(Nichols 1994: 21–22)	Chechen	sg	w	j	j	d	b	b
(Nichols 1994: 21–22)	Chechen	pl	b	d	j	d	b	d

As mentioned earlier, the attribution of the masculine and feminine noun classes is very transparent. The animate class consists of animals, both large and small (the only exception are insects, which in some cases are grouped with inanimates). The motivation for assigning any of the inanimate classes remains a mystery. Not even all insects belong to the same inanimate class: in Rikvani, spiders belong to the animate class, while butterflies belong to the third inanimate class. In addition, *aqrab* 'scorpion', which is a loan from Avar, also belongs to the animate class, while the native

word *hortʃ'oritʃ'in* 'scorpion, belongs to the third inanimate class. A possible explanation for *aqrab* belonging to inanimate 1 is analogy with Avar - in Avar there is only one neuter class, which comprises animals and objects, and corresponds to the class marker *b-*, parallel to animate and inanimate 1 in the Andi dialects. In Figure 3. the number of classes in Andi and neighbouring villages is presented.

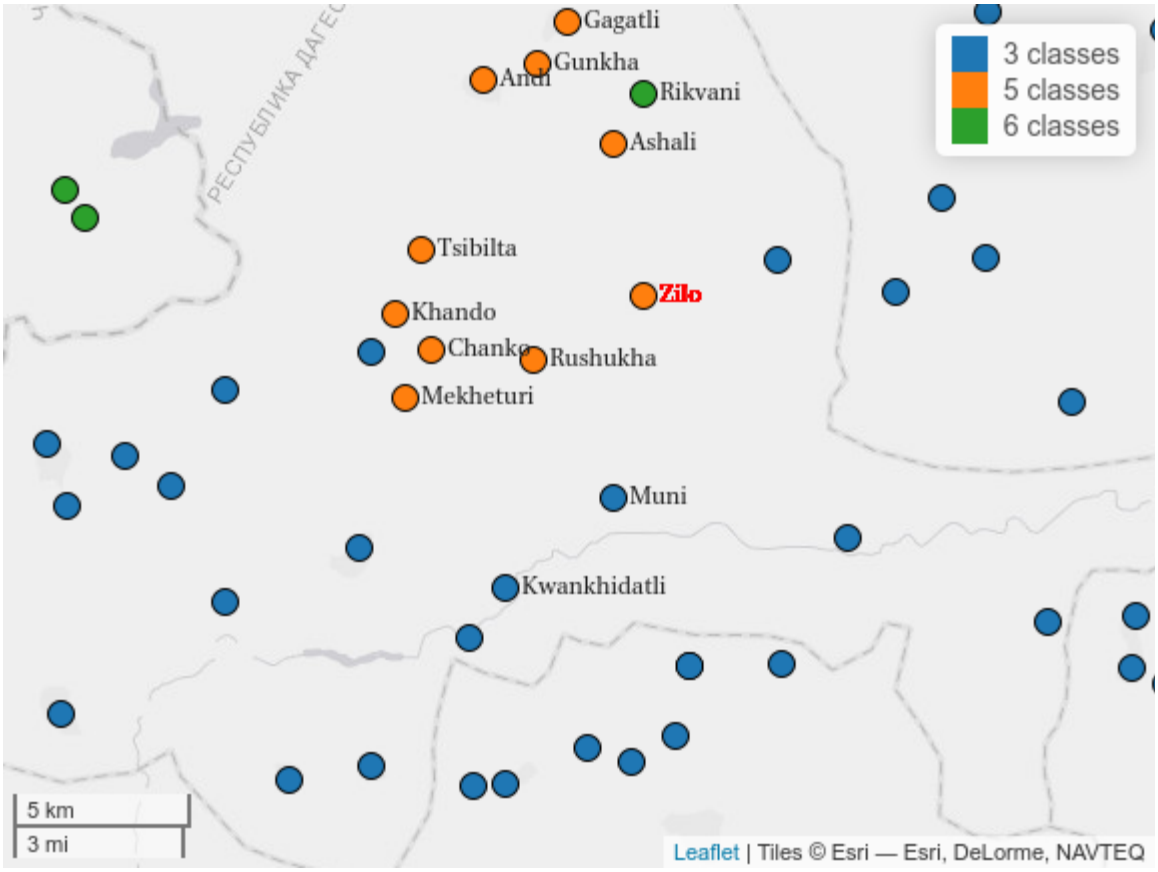

Figure 3. Number of classes in Andi and their neighbours.

3.1 Inanimate classes

In the descriptive grammar by (Suleymanov 1957) (Rikvani dialect), a dictionary of 1236 nouns is provided, grouped by semantic category and noun class. The vast majority of these nouns belongs to the inanimate 1 class (678). There are 53 masculine nouns, 38 feminine, and 46 ambiguous human nouns. A total of 29 nouns is classified as ambiguous between masculine, feminine and animate (e.g. *motʃi* 'child'). The animate class consists of 109 nouns. There are 276 nouns which belong to inanimate 2, and the third inanimate class contains just seven insects. As the first inanimate class comprises half of the noun vocabulary, it covers a large number of rather general semantic categories, including body parts, plants, household objects and abstract concepts (Suleymanov 1957: 138–151). Unfortunately, the second inanimate class seems to cover largely the same categories, just different parts of it. For example, *relʕa* 'hand' belongs to the first inanimate class, while *вазu* 'arm' belongs to the second. Few of the categories are exclusive as a whole to either inanimate 1 or inanimate 2, with the exception of a few highly specialized categories belonging to inanimate 2: religious objects and moral concepts (e.g. prayer, tax, amulet); paper objects and documents (book, letter, newspaper, various legal

documents); concepts of time (hour, day, night). Curiously, female clothing and jewelry is mostly inanimate 2, while male clothing predominantly belongs to inanimate 1.The situation is a bit different in Gagatli, as described in (Salimov 2010 (1968): 48). In Gagatli, the feminine class contains a handful of abstract concepts that belong to inanimate 2 in Rikvani (e.g. evening). Otherwise, the description of the masculine and animate classes, as well as that of the inanimate 1 and 2 classes is highly similar. The insects distinguished by the third inanimate class in Rikvani are grouped with inanimate 2 in Gagatli. For both dialects liquids and water reservoirs tend to be assigned the inanimate 2 class. For Zilo this largely hold, although *inχi* 'large river' belongs to inanimate 1.

3.2 Origins of inanimate 2

A few remarks on diachrony are in order here. As Table 2 illustrates, the prefixal marker of inanimate plurals is similar to the inanimate 2 marker in Andi. Three-class systems are considered reduced systems originating in a four or five class system that is more similar to what we find in contemporary Andi (Alekseev and Atayev 1997). As argued in (Daniel 2018) there is some evidence that across the family, the second inanimate class (in cases where there is such a distinction) and similar-looking plural markers, ultimately originate in a distinct class containing mass nouns, such as liquids, pluralia and singularia tantum and abstract nouns that are "number neutral". If that is the case, the inanimate 2 has extended its terrain considerably in contemporary Andi. It contains some mass nouns and singularia tantum like *ilajma* 'parents' (literally a contraction of 'mother' (*ila*) and 'father' (*ima*), which agrees with inanimate 2 singular), but also many countable objects. In addition, mass nouns are not restricted to inanimate 2, for example *zar* 'ice' and *k'ːõj* 'smoke' belong to inanimate 1 at least in Rikvani. Thus, evidence from Andi does not necessarily refute the mass noun hypothesis, but it cannot support it either, as we have no knowledge of how Andi noun classes have evolved. Moreover because we have not yet succeeded in establishing semantic regularities for the synchronic level.

3.3 Loan words

Based on the lists by Sulejmanov and Salimov, we can conclude that there is no distinct pattern for loan words from certain languages, or loan words in general as a group, as we find borrowings in each class: Russian names of professions in the masculine and feminine classes, names of religious entities which are usually identified as male and borrowed from Arabic are in the masculine class, and various objects borrowed from different languages belong to inanimate 1 or 2. Words referring to modern technology are similarly distributed among the inanimate classes in an unpredictable way. For example *televizor* 'television' is inanimate 1, while *telefon* 'telephone' belongs to the inanimate 2 class. As we will show in the following section, this lack of semantic transparency and regularity does not hinder speakers in being surprisingly consistent with regards to the class they assign to nouns in general, and new nouns in particular.

4. Experimental data and analysis

In order to evaluate how consistent speakers are in assigning a noun class to certain nouns, we created a list consisting of 25 native and 25 borrowed words for each of the inanimate classes (100 words in total) and some members of the inanimate 3 class from Rikvani (four in total). We included native words in order to detect a difference in variation across speakers between words that are part of the core lexicon and words that are relatively new and obviously come from a different language. When

the list was compiled, we arranged the words in a quasi-arbitrary order; we tried to avoid offering the speakers large clusters of borrowings and too much repetition of words belonging to the same class, in order to avoid priming. The complete list of 114 lexical items as well as all R code[3] for this article are available online: https://github.com/LingConLab/2018_Zilo_classes_article. For each of the words, we asked speakers to translate the phrase "my X", because the possessive construction in Andi consists of a personal pronoun with a class suffix that agrees with the possessee, for example: *di-w ima* (I-m father) 'my father'. We conducted the experiment with sixteen different speakers, evenly divided among two age groups: younger (13-32) vs. older (34-60). For each age group, we interviewed four female and four male speakers. All of the data were collected during fieldwork in Zilo in August of 2017.

Two measures were used to determine the uniformity of the speakers' answers: the percentage of absolute agreement and Fleiss' Kappa. The characteristics of the obtained values of the Fleis' kappa were taken from (Hallgren 2012). The results are shown in Table 3:

Table 3. Pure agreement percentages and Fleiss' kappa for all speakers

	% abs. agr.	Fleiss' Kappa	z-score	characteristic
all observations	74.5	0.849	95.7	Almost perfect agreement
native words	84.2	0.937	77.5	Almost perfect agreement
borrowings	63.3	0.732	56.1	Substantial agreement

As Table 3 shows, the results of the experiment showed a fairly high inter speaker agreement for the class assignment task. The agreement percentage for the group of borrowed words is lower (substantial agreement vs. almost perfect agreement). For the native words agreement is quite high, but not ideal. This shows that variability in class assignment is observed in all parts of the vocabulary, although it is much higher for borrowings.

The results of a hierarchical clustering of our data are presented in Figure 3:

3 All statistical tests and visualizations were done in R (R Core Team 2018), using the following packages: *ape* (Paradis, Claude, and Strimmer 2004), *ggplot2* (Wickham 2009), *irr* (Gamer, Lemon, and Singh 2012) and *lingtypology* (Moroz 2017).

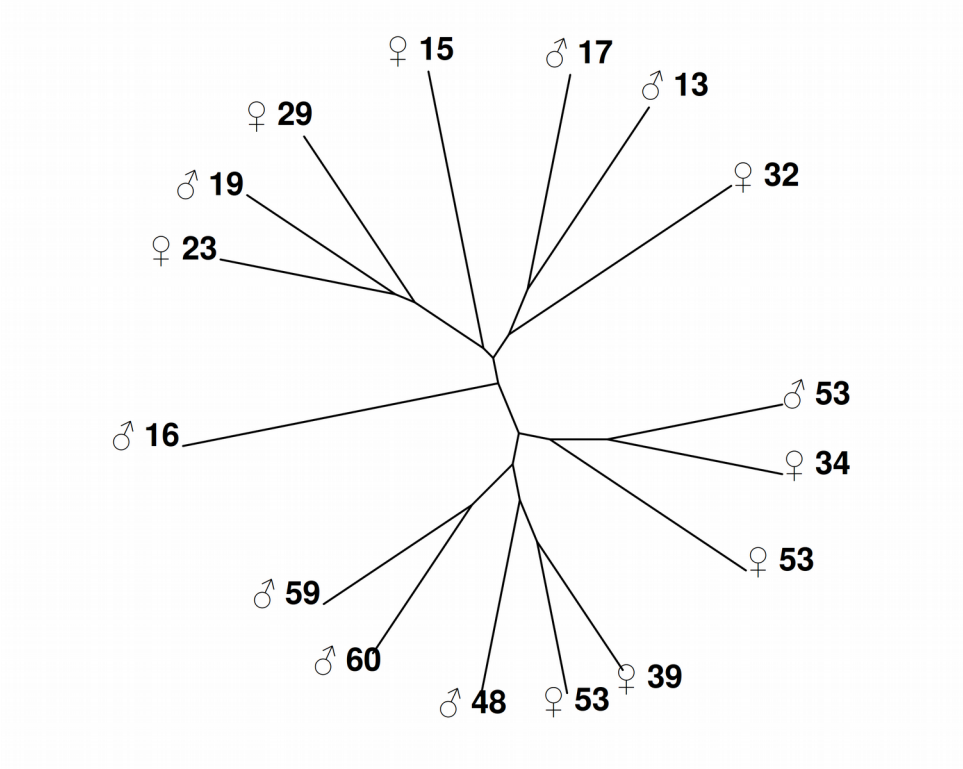

Figure 4. Dendrogram showing the clusterization of speakers.

On the dendrogram in Figure 3, three different clusters can be distinguished, which correspond to age groups:

- older group from 34 to 60;
- younger group from 13 to 32;
- one sixteen year old female speaker forms a separate cluster.

Since hierarchical clustering demonstrates the significance of age groups, we recalculated all of the obtained measures for the two age groups:

Table 4. Pure agreement percentages and Fleiss' kappas for different age groups

	age gr.	% abs. agr.	Fleiss' Kappa	z-score	characteristic
all	older	85.85	0.887	48.3	Almost perfect agreement
all	younger	79.25	0.845	46.1	Substantial agreement
native	older	92.98	0.943	37.7	Almost perfect agreement
native	younger	87.72	0.928	37.1	Almost perfect agreement
loan	older	77.55	0.804	29.8	Substantial agreement
loan	younger	69.39	0.741	27.4	Substantial agreement

As we see from Table 4, younger speakers have lower values for both the pure agreement percentages and Fleiss' kappas in each group of words. There are two possible scenarios for these values: 1) variability concerns the same lexemes in the different age groups, but that the level of

variability is higher among younger speakers; 2) the set of variable lexemes is different for the two groups. Therefore, we decided to visualize which lexemes are most variable for each group. First, we calculated variability rate for each word in each group. Here are three examples:

- the word *ingur* 'window' was attributed inanimate 1 class by all speakers, so its variability rate will be
$$\frac{inanimate_1}{inanimate_1+inanimate_2}=\frac{16}{8+8}=1$$
- the word *haq'u* 'room' was attributed inanimate 1 class by all speakers, so its variability rate will be
$$\frac{inanimate_1}{inanimate_1+inanimate_2}=\frac{0}{8+8}=0$$
- the word *odoruk'a* 'butterfly' was attributed inanimate 1 class by older speakers four times, and by younger speakers five times, so its variability rate will be different in different groups:

$$\frac{inanimate_1}{inanimate_1+inanimate_2}=\frac{4}{4+4}=0.5 \qquad \frac{inanimate_1}{inanimate_1+inanimate_2}=\frac{5}{5+3}=0.625$$

In Figure 4, the variability rate of younger speakers is plotted against the variability rate of older speakers.

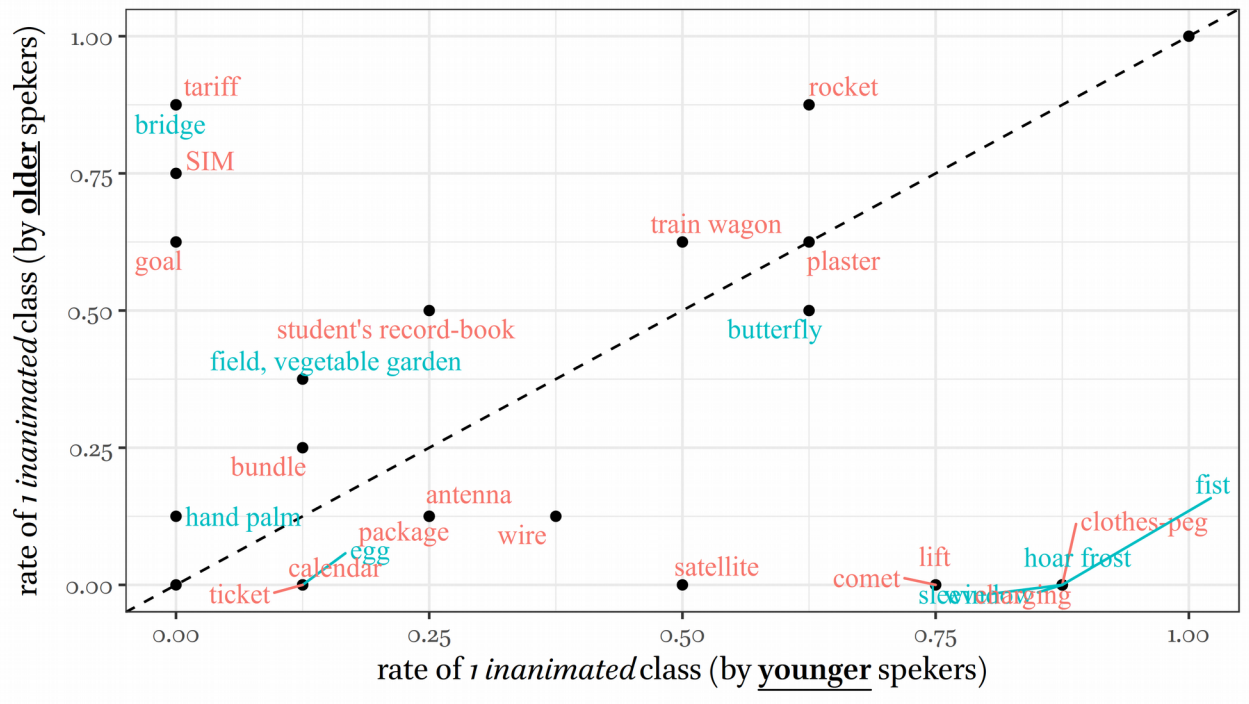

Figure 5. Distribution of the most variable words by different age group (native lexemes are in blue).

Lexems that were affiliated with the first inanimate class will have a variability rate of 1 for both groups, so it will appear in the top right corner. Lexems that were affiliated with the second inanimate class will have a variability rate of 0 for both groups, so it will appear in the bottom left corner. In case

a lexeme was affiliated with one inanimate class by one group and with another by the other group, it would appear in one of the other corners, but there are no such cases in our data. Since cases with no variability are not interesting for us, these dots are not labeled in the graph. All other points in Figure 4 show the variability preferences for each age group. Lexemes that appear under the dashed line (with the formula $y = x$) tend to be more variable among the older speakers, and words that appear under the line tend to be more variable among younger speakers. So from Figure 4 we clearly see that the variability of some lexems (only 27 of 106) is shared between both age groups (e.g. *plastirʲ* 'plaster'), but there are also some lexems that are variable in one group and not variable in the other (e.g. *sputnʲik* 'satellite' or *gol* 'goal').

5. Conclusion

The Upper Andi dialects have an unusually large number of noun classes in comparison with neighbouring languages and other Nakh-Daghestanian languages. This is caused by the distinction of multiple classes for inanimate objects. While the assignment of the masculine, feminine and animate classes is clearly semantically motivated, we could discern no obvious pattern for the assignment of a particular inanimate class. Therefore, we decided to conduct an experiment testing the class attribution of 114 words referring to inanimate objects (and a few insects that are usually categorized with inanimate objects. The speakers of the Zilo dialect of Andi showed high levels of consent. As our experiment showed, variation occurs both in the native and the borrowed vocabulary, though for borrowings it is much higher. A hierarchical clustering of the experimental results divided the group of speakers into older and younger speakers. Variation is higher among the younger speakers. In addition, variation in class asignment is not homogeneous among the age groups. Variable lexemes can be divided into those that are shared by the two age groups, and those which are unique for each age group.

Bibliography

Alekseev, M. E. 1999. "Andijskie Jazyki." In *Jazyki Mira. Kavkazskie Jazyki*, edited by M. E. Alekseev, S. A. Starostin, G. A. Klimov, and J. G. Testelets, 220–28. Moscow: Academia.

Alekseev, M. E., and B. M. Atayev. 1997. *Avarskij Jazyk* [*the Avar Language*]. Moscow: Academia.

Daniel, M. 2018. *Mass and Class. Interaction of Number, Class and Countability in* (*Several*) *East Caucasian Languages, with Special Focus on Archi and Mehweb*. (Handout).

Dobrushina, N.R. 2011. "Mnogojazyčie V Dagestane Konca Xix-Načala Xxi Veka: Popytka Količestvennoj Ocenki [Multilingualism in Daghestan from the End of the Xix Century - Beginning of the Xxi Century: An Attempt at Quantitative Analysis]." *Voprosy Jazykoznanie*, no. 4: 61–80.

Gagliardi, Annie, and Jeffrey Lidz. 2014. "Statistical Insensitivity in the Acquisition of Tsez Noun Classes." *Language*, no. 90: 58–89.

Gamer, Matthias, Jim Lemon, and Ian Fellows Puspendra Singh. 2012. *Irr: Various Coefficients of Interrater Reliability and Agreement*. https://CRAN.R-project.org/package=irr.

Gudava, T. E. 1962. *Botlixuri Ena* [*the Botlikh Language*]. Tbilisi: AN GSSR.

Hallgren, K. A. 2012. "Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial." *Tutorials in Quantitative Methods for Psychology* 8 (1). NIH Public Access: 23.

Magomedbekova, Z. M. 1971. *Karatinskij Jazyk: Grammatičeskij Analiz, Teksty, Slovar'* [*the Karata Language: Grammatical Analysis, Texts, Dictionary*]. Tbilisi: Metsniereba.

Moroz, George. 2017. *Lingtypology: Easy Mapping for Linguistic Typology*. https://CRAN.R-project.org/package=lingtypology.

Nichols, J. 1994. “Chechen.” In *The Indigenous Languages of the Caucasus, Vol. 4: The North East Caucasian Languages, Part 2*, edited by R. Smeets, 2–77. Delmar, NY: Caravan Books.

Paradis, E., J. Claude, and K. Strimmer. 2004. “APE: Analyses of Phylogenetics and Evolution in R Language.” *Bioinformatics* 20: 289–90.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Salimov, X. S. 2010 (1968). *Gagatlinskij Govor Andijskogo Jazyka* [*the Gagatli Dialect of the Andi Language*]. Makhachkala: IJaLI.

Schulze, Wolfgang. 2013. “Historical and Areal Aspects of Natural Resource Terminology in East Caucasian Languages.” *IRAN AND THE CAUCASUS* 17 (3). Brill Academic Publishers, Leiden, Netherlands: 295–320.

Stolz, Christel. 2009. “A different kind of gender problem: Maltese loanword gender from a typological perspective.” In *Introducing Maltese Linguistics.*, edited by Bernard Comrie, Ray Fabri, Elizabeth Hume, Manwel Mifsud, Thomas Stolz, and Martine Vanhove, 321–54. John Benjamins.

Suleymanov, J. G. 1957. “Grammatičeskij Očerk Andijskogo Jazyka (Po Dannim Govora S. Rikvani) [Grammar Sketch of the Andi Language (Based on Material from the Dialect of the Village Rikvani)].” PhD thesis, Institut Jazykoznania AN SSSR.

Tsertsvadze, I. I. 1965. *Andiuri Ena* [*the Andi Language*]. Tbilisi: Metsniereba.

Wickham, Hadley. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. http://ggplot2.org.