Tolerance of underspecification
Evidence from Karachay-Balkar*

Pavel Koval & Alëna Aksenova
Lomonosov Moscow State University

Abstract. In this paper we discuss the microhistoric variation in the Cherek dialect of the Karachay-Balkar language. We argue that a number of differences between the language of two generations is due to a tiny structural change on the interface between syntax and PF. Two types of evidence are discussed: possessive constructions and number mismatch on bound pronouns.

Keywords. microhistoric variation, possessive constructions, number match, \( \varphi \)-features, underspecification, Turkic languages

1. Setting the scene

The question of how language changes has been around for some time (King 1969, Campbell 1998). Many studies of specific instances of language change offer a perspective of at least 2-3 centuries. The great part suffers from the same weakness: lack of negative evidence. It is impossible to say for sure which structures were impossible for the ancients, the question itself was formulated less than a century ago (Chomsky 1957). In this squib we offer a case study of a microhistorical structural change between two generations of speakers. It is the minimal "distance" at which a difference can be observed. We present novel data from an understudied minor Turkic language, where the change is rather drastic. We argue that a number of differences can be reduced to a structural change at the interface between syntax and PF.

In the Cherek dialect of the Karachay-Balkar language (Turkic, Kypchak) there is a bundle of differences between the language of the old generation (>45 y.o., \( G_1 \)) and of the young one (<30 y.o., \( G_2 \)). Due to space limit, we discuss in detail two differences coming from the nominal structures: syntax of possessive constructions and number agreement on bound pronouns. We argue that the syntactic structures remain the same in \( G_1 \) and \( G_2 \). What has changed is the rule of Spell-Out of underspecified structures. In a nutshell this change can be stated as follows:

---

* Many thanks to the audience and organizers of SinFonIJA VII and two anonymous reviewers for helpful discussions and insightful comments. Thanks to all participants of the Balkar linguistic expedition (aul Ullu Malqar), and especially to its organizers Sergei Tatevosov, Pavel Graschenkov, and Ekaterina Lyutikova. We gratefully acknowledge support from RFBR (grant # 13-06-00884a). All the errors are the authors' own.
1. Tolerate nominal structures with underspecified \( \varphi \)-features for Spell-Out. \((G_1)\)
2. Do not tolerate nominal structures with underspecified \( \varphi \)-features, leave it unpronounced. \((G_2)\)

We start with the discussion of the phenomena, then we show how our proposal can be applied to the facts, and we end with speculations about the actuation problem.

2. Syntax of possessive constructions

Both the system of \(G_1\) and \(G_2\) have two possessive constructions (PC). The first is the construction with a referential possessor (PC I, see (1)); the other features a non-referential (i.e. generic) possessor (PC II, see (2)). PC I looks the same way in both \(G_1\) and \(G_2\), while PC II differs, cf. (2) and (3).

(1) Alim-ni surat-i
    Alim-GEN picture-3
    'Alim's photo' \((G_1 \& G_2)\)

(2) sabij bolnica-si
    Child hospital-3
    'children's hospital' \((G_1)\)

(3) sabij bolnica
    Child hospital
    'children's hospital' \((G_2)\)

For \(G_2\) presence of the referential possessor marker implies presence of the referential possessor in the interpretation, cf. (4) where the first interpretation is impossible for \(G_2\).

(4) tiširiu zijriH-i
    Woman dress-3
    'a women's dress' \(^{ok}G_1 \ast G_2\)
    'the women's dress owned by someone' \(^{ok}G_1 \ast G_2\)

In both \(G_1\) and \(G_2\) the referential possessive marker agrees with the possessor in person (5), while the non-referential one in \(G_1\) always stands in the form of the 3\(^{rd}\) person.

(5) meni kitab-im
    my.GEN book-1
    'a women's dress' \(^{ok}G_1 \ast G_2\)
    'the women's dress owned by someone' \(^{ok}G_1 \ast G_2\)
Adopting the standard view on the organization of person features\(^1\) (Zwicky 1977, Noyer 1992 \textit{i.e.}), this can be seen as a case of underspecification. In other words, \(G_2\) lacks the underspecified marker of possession.

It should also be noted that the plural marker always precedes the possessive marker in both \(G_1\) and \(G_2\) (6-7).

\begin{align*}
(6) & \text{ qiz-ni } \quad \text{ginži-ler-i} \\
& \quad \text{girl-GEN} \quad \text{doll-PL-3} \\
& \quad \text{the girl's dolls'}
\end{align*}

\begin{align*}
(7) & \text{ *qiz-ni } \quad \text{ginži-si-ler} \\
& \quad \text{girl-GEN} \quad \text{doll-3-PL}
\end{align*}

We adopt Grashchenkov (2007)'s view on the syntax of possessive constructions in Turkic. According to this view, within the noun phrase two projections take care of possessors: Spec nP licenses the non-referential possessor and Spec DP assigns genitive to the referential possessor (cf. Split-NP analysis in Radford (2000) \textit{i.e.}). According to Split-NP analysis, the number projection NumP mediates nP and DP. Heads n, Num, and D bear \([\text{POS}], [\text{NUM}], \text{ and } [\text{PERS}]\) features respectively. This can be illustrated with the following tree:

Fig. 1

\begin{center}
\begin{tikzpicture}

\begin{scope}[level distance=1.5cm, level 1/.style={sibling distance=3.5cm, level distance=1.5cm}]

\node{DP}
\node (PCII) [below] {PC II \textbf{[+GEN]}}
\node (NumP) [below] {NumP \textbf{[+PERS]}}
\node (D) [below] {D \textbf{[+PERS]}}
\node (nP) [below] {nP \textbf{[+NUM]}}
\node (Num) [below] {Num \textbf{[+NUM]}}
\node (PCI) [below] {PC I}.
\end{scope}

\begin{scope}[level distance=1.5cm, level 2/.style={sibling distance=2cm}]

\node{} child {node {NP \textbf{[+POSS]}}
child {node {n}}}
child {node {Num}}
child {node {D}}
child {node {NumP}}
child {node {nP}}
child {node {PCII}}
child {node {PCI}}
\end{scope}

\end{tikzpicture}
\end{center}

In \(G_1\) the possessive marker can be sent to the Spell-Out domain with the underspecified \([\text{PERS}]\) feature\(^2\), i.e. at the level of Num (or n, depends on our rules of linearization), while in \(G_2\) this is no more the option.

\footnotesize
\(^1\) In what follows all the discussed features should be read as shorthands for bundles of features, e.g. a couple of binary \([\text{speaker}]\) and \([\text{hearer}]\) features forms \([\text{pers}]\) feature.

\(^2\) The relation between person features and the notion of underspecification merits discussion, but space precludes it, cf. Rooryck 1994, Bobaljik et al. 2011, Müller 2011.
3. Number features on bound pronouns

$G_1$ licenses the use of a 3$^{rd}$ person singular pronominal with the semantically plural antecedent (8), while $G_2$ requires full feature match between an antecedent and its anaphor.

(8) [eki džaščik], [an-ni, egeč-i-ni] sūj-e.
   two boy (s)he-GEN sister-3-ACC love-PRS
   Two boys love their sister. ($G_1$)

Speakers of $G_2$ accept examples like (8) only without coindexation of the pronoun and the noun phrase. For $G_2$ the only possible way to coindex them is to use morphologically plural pronoun (9), for $G_1$ this structure is acceptable, but dispreferred.

(9) [eki džaščik], [an-nar-ni, egeč-i-ni] sūj-e.
   two boy (s)he-PL-GEN sister-3-ACC love-PRS
   Two boys love their sister. ($G_1$ & $G_2$)

As expected, both $G_1$ and $G_2$ ban the match of a morphologically plural anaphor and a singular antecedent (10). This suggests that examples like (8) are better to be seen as involving underspecification.

(10) *džaščik, [an-nar-ni, egeč-i-ni] sūj-e.
    boy (s)he-PL-GEN sister-3-ACC love-PRS

What we argue for is that the pronominal in (8) for $G_1$ is underspecified for number (cf. similar ideas proposed for different languages in van Gelderen 1992, Burzio 1989 i.m.a.).

One way to see this is to assume that the number feature here is semantically inert (Partee 1989, Kratzer 1998, 2009, Heim 2008). In line with a minimal pronoun account (Kratzer ibid.), we have to claim that such features are purely morphological and that they are transmitted from a binder to a pronoun it binds via a PF operation (Kratzer ibid., Heim ibid.).

The alternative account (Sudo 2012, Nikolaeva 2014) puts most of the burden on syntax. Under this view binder and bindee share their $\varphi$-features via coindexation in the syntactic module.

The proposed shift between two rules is licit under both accounts. The second theory is perfectly compatible with it, while for the minimal pronoun theory our rule should follow the proposed PF operation.

4. Conclusion

We have seen two differences between two systems used by two generations of speakers of the Cherek dialect of the Karachay-Balkar language. The first difference is that the possessive marker which originally sufficed for both referential and non-referential (generic) possessors became 'reinterpreted' as a marker of referential possessor. The second difference is that $G_2$ speakers allow coindexation for two
referents if all their φ-features match, while \( G_1 \) tolerate underspecification in number. The proposed solution explains both of these differences as residual effects of a general shift to more marked structures. The understandable weakness of this proposal is its empirical evidence, since the proposed general shift should have generated much more differences. Balkar is an understudied language, and due to time limits in expedition and space limits for this paper we had to discuss in detail two differences of those discovered. All we can say is that at least we are unaware of any contradicting cases.

One could think about various extraneous triggers for such change, it is hard to determine the real reasons even in this particular case. The question we would like to address instead is how this change can be stated in terms of marked and unmarked values of the parameter, how this change affects the language itself, and what could be the motivation for it.

One possible approach to account for this shift is the Subset Principle (Berwick 1985). According to this principle, the child first chooses the smallest value of a parameter, and goes beyond when she encounters the positive evidence for a larger grammar. Since the grammar of \( G_2 \) is smaller and more marked than the one of \( G_1 \) while sets of meanings that can be expressed with these grammars are equal and bijective, the \( G_2 \) system is more economical and therefore it escapes penalty for disobeying the Subset Principle. The important question is why the \( G_1 \) grammar existed at all, but an answer to this question is far beyond the scope of this paper.

5. List of glosses

1 — possessive marker of the first person;
3 — possessive marker of the third person;
ACC — accusative case;
GEN — genitive case;
NUM — number feature;
PERS — person feature;
PL — plural;
POSS — possessive feature;
PRS — present tense.

References


