Babby's puzzle: syntax or morphology?

Asya Pereltsvaig (Stanford University)

1. Introduction. Babby (1987, and elsewhere) drew a distinction between the heterogeneous and the homogeneous case distributions in Russian quantified noun phrases. Since then, the question arose as to whether these morphosyntactic facts should be properly accounted for in syntax or in morphology. Bailyn (2004) proposed a syntactic account of the hetero-/homogeneous case distribution, according to which the numeral is merged in a Spec-QP with the heterogeneous case distribution and as a head Q° with the homogeneous case distribution; moreover, according to his analysis, with the heterogeneous case distribution the null Q° assigns/checks GEN case on its complement. However, this syntactic account leaves open several issues concerning the form of (i) the noun in certain quantified noun phrases, (ii) attributive adjectives in certain quantified noun phrases, (iii) prequantifiers, as well as (iv) the alternation between regular and collective numerals. This paper's contribution is two-fold: first, it uses corpus study and large scale native speaker surveys to collect data on the grammaticality and use of morphological forms in quantified noun phrases in Russian. Second, it lays out a uniform account of a range of "unexpected" and "messy" facts, placing the explanatory burden in morphology and the lexicon.

2. Data. Syntactic analyses, such as Bailyn's (2004) or Ionin and Matushansky's (2004, 2006), maintain that with the heterogeneous case distribution the head Q° (null or overt, depending on the analysis), assigns/checks genitive case to its complement, and the contrast between GEN.PL with "higher numerals" (5+) and GEN.SG with paucal numerals (2-4) is ascribed to the selectional properties of the two types of numerals. Such an analysis fails to account for the following **four facts** involving the heterogeneous case pattern: **(a)** special count forms of certain nouns, distinct from GEN.PL or GEN.SG, (**1**); **(b)** the ungrammaticality of GEN.SG forms of certain classes of feminine nouns (deadjectival nouns and last names in *-ova*) with paucal numerals – NOM.PL or GEN.PL forms are used instead, (**2**); **(c)** similarly, with paucal numerals and feminine nouns, the form of the attributive adjective is not GEN.SG but NOM.PL or GEN.PL, as in (**3**) – this is true regardless of whether or not a given noun has homophonous GEN.SG and NOM.PL forms; **(d)** the form of the prequantifier is typically GEN.PL, regardless of the morphological form of the noun, as in (**4**). Furthermore, existing literature (descriptive, prescriptive and theoretical alike) provides **no coherent description** of **(a)** the variation between NOM.PL and GEN.PL forms in structures like (**2**)-(**3**), and **(b)** the alternation between regular and collective numerals in (**5**). **Our study** involving both corpus study and a large scale speaker survey (over 60 speakers have responded to date) shows that although for some speakers the GEN.PL form is preferred in (**2**)-(**3**), for the majority of the speakers the NOM.PL form is the preferred one, regardless of whether the noun has homophonous GEN.SG and NOM.PL forms or not, as in (**3a**) vs. (**3b**) (contrary to claims in Graudina et al. 1976 and elsewhere). In addition, we show that – again, contrary to claims in the literature – collective numerals are acceptable with both masculine and feminine nouns. Moreover, our study also uncovered the possibility of NOM.PL form for the prequantifier, esp. in the presence of a NOM.PL attributive adjective, as in (**6**), previously unmentioned in the literature. Partial results of native speaker survey are diagrammed in (**7**).

3. Analysis. We work out a detailed analysis of the abovementioned morphosyntactic facts based on three core ideas. The first (not particularly novel or controversial) idea is the distinction between abstract syntactic Cases and their morphological realization: a given morphological case (m-case) may spell-out a number of distinct syntactic Cases and conversely a given syntactic Case may be spelled-out by a number of distinct m-cases. We take all elements in a quantified noun phrase to be syntactically specified as QUANT (Quantificational Case; cf. Bailyn 2004, Rappaport 2004), which may be spelled out as either nominative or genitive m-case. Second, Lexical Insertion rules operate on bundles of features, including categorial features, Case, number, gender, animacy (and for nouns, declension class); this is also taken to account for accusative syncretism patterns. Third, we argue that all elements of a quantified noun phrase are fully specified for Case, gender and number. In particular, we propose that numerals are inherently specified for number (PL for "higher numerals" and PAUC for 2-4; we take PAUC to be a number rather than a Case, contrary to Rappaport 2004), while nouns are not inherently specified for number (as a result bare NPs are spelled out as PL in Russian but as SG in English; cf. Pesetsky 2007, and (**8**) below). These hypotheses allow us to account for various morphological forms, variability of usage and language change.

% marks structure grammatical for some but not all speakers

(1) *pjat' čelovek* / **ljudej*, *dva časá* /**časa*
five people.COUNT / *people.GEN.PL two hour.COUNT /*hour.GEN.SG

(2) a. *dve gorničnye* /% *gorničnyx* / **gorničnoj*
two maids.NOM.PL / maids.GEN.PL/ *maid.GEN.SG

b. *tri Ivanovy* / % *Ivanovyx* / **Ivanovoj*
three Ivanova.NOM.PL / Ivanova.GEN.PL / *Ivanova.GEN.SG

(3) a. homophonous GEN.SG and NOM.PL (both *nóvosti*)

dve { *xorošie* / *xorošix* / **xorošej* } *novosti*
two good.NOM.PL / good.GEN.PL / *good.GEN.SG news

b. non- homophonous GEN.SG and NOM.PL (*gorý* vs. *góry*)

dve { *vysokie* / *vysokix* / **vysokoj* } *gory*
two tall.NOM.PL / tall.GEN.PL / *tall.GEN.SG mountains

(4) a. *celyx dve gorý*
whole.GEN.PL two mountain.GEN.SG

b. *celyx polčasá*
whole.GEN.PL half-hour.COUNT

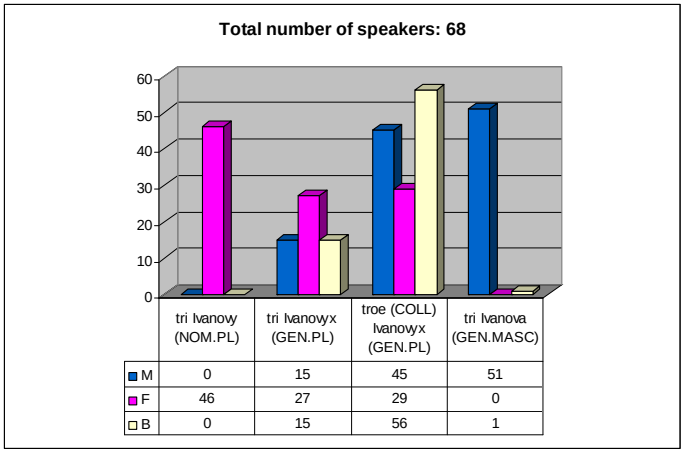
(5) a. *tri* { *mal'čika* / *devočki* / **brjuk*}
three boy.GEN.SG / girl.{GEN.SG?/NOM.PL?} / *pants.GEN.PL

b. *troe* { *mal'čikov* / *%devoček* / *brjuk*}
three.COLL boys.GEN.PL / girls.GEN.PL / pants.GEN.PL

(6) a. *%celye dve novosti*
whole.NOM.PL two news.{GEN.SG?/NOM.PL?}

b. *%celye dve xorošie novosti*
whole.NOM.PL two good.NOM.PL news.{GEN.SG?/NOM.PL?}

(7) Speakers accepting each of the four structures as referring to a group of three boys (M), three girls (F) or a mixed group (B)



(8) RU: *remont kompjuterov* vs. EN: *computer repair*
repair computers.GEN.PL computer.SG repair