

# АНАЛИЗ РЕЧЕВЫХ ОШИБОК ПРИ ПРЕДИКАТИВНОМ СОГЛАСОВАНИИ В РУССКОМ ЯЗЫКЕ: ЭФФЕКТ РОДА ГЛАВНОГО ИМЕНИ

## SUBJECT-VERB AGREEMENT ERRORS IN RUSSIAN: HEAD NOUN GENDER EFFECT

*И.С. Янович (Igor\_Y@abbyy.com)*  
МГУ им. М.В. Ломоносова, ABBYY Software House

*О.В. Фёдорова (olga.fedorova@msu.ru)*  
МГУ им. М.В. Ломоносова, Москва

В работе приводятся новые данные, показывающие наличие влияния грамматического рода главного имени на процессы предикативного согласования по числу в русском языке. Это влияние объясняется различной степенью маркированности разных грамматических родов в русском языке.

### 1. Введение

Во многих современных психолингвистических исследованиях частная проблема грамматического согласования связывается с глобальной проблемой соотношения языка и мышления. Современные последователи гипотезы Сэпира-Уорфа и Вюрцбургской школы сохраняют все ту же поляризацию взглядов: в одних теориях (см., например (Bates & MacWhinney 1989)) язык представляется в виде неотъемлемой части общего когнитивного механизма человека, в других же он рассматривается как самостоятельно функционирующая подсистема, излированная от перцептивной, когнитивной и моторной систем (Fodor 1983). Из двух вышеизложенных позиций вытекают две полярные гипотезы, стремящиеся объяснить, как на самом деле происходит процесс предикативного согласования по числу между подлежащим и сказуемым. Согласно первой из них, получившей название **понятийной**, при предикативном согласовании множественность последовательно кодируется сначала в именной группе (ИГ), а затем в глагольной. Таким образом, предикативное согласование является следствием раздельной лексикализации отдельных частей одного и того же высказывания; каждая из этих частей сообщает что-то свое о множественности события. Согласно второй, **флективной**, гипотезе, грамматическое число подлежащего выбирается исходя из понятийного числа, а затем, в процессе порождения высказывания, это грамматическое число копируется в глаголе.

### 2. История изучения согласования подлежащего и сказуемого по числу

Вопрос о процессе предикативного согласования – один из самых обсуждаемых в психолингвистической периодике последних лет: с момента опубликования пионерской работы (Bock & Miller 1991) появилось уже более полусотни работ, непосредственно посвященных этой тематике. Авторы первых подобных исследований изучали только процессы **порождения** речи исключительно в **английском** языке. В середине 90-х годов ситуация стала меняться: с одной стороны, появились эксперименты на материале итальянского (Vigliocco et al. 1995), испанского (Vigliocco et al. 1996b), французского и голландского (Vigliocco et al. 1996a), а также русского (Nicol & Wilson 1999) языков. С другой стороны, исследователи стали интересоваться проблемами согласования при **понимании** высказывания, используя методики чтения с регулировкой скорости и записи движения глаз.

Процесс предикативного согласования при порождении речи обычно изучают, анализируя речевые ошибки, возникающие в ходе эксперимента на **заканчивание предложений**: испытуемый читает начало предложения, состоящее из сложной ИГ с разными комбинациями по числу главного (Гл) и локального (Лок) имен, например, *The keys to the cabinet*, и, повторив эту преамбулу, заканчивает предложение - *The keys to the cabinet were lost*.

Давно известно, что распределение **речевых ошибок** в подобных экспериментах несимметрично: их больше всего, когда ГлИГ стоит в единственном числе (Sg), а ЛокИГ – во множественном (Pl) (*\*The key to the cabinets were lost*). Самое распространенное объяснение этого факта таково: имена существительные по умолчанию единичны, и в случае множественности их необходимо как-то специально маркировать. Поэтому ГлИГ, стоящая в Sg, более уязвима и подвержена интерференции со стороны ЛокИГ, стоящей в Pl. Предложения, в которых ГлИГ стоит в Pl, а ЛокИГ – в Sg, следовательно, более устойчивы по двум причинам: с одной стороны,

ГЛИГ уже маркирована, а с другой – ЛокИГ, имея Sg по умолчанию, не обладает большой силой притяжения (Eberhard 1997).

Одним из самых серьезных и одновременно дискуссионных вопросов в данной проблематике является вопрос о том, на каком уровне порождения высказывания происходит процесс согласования. Большинство современных моделей речепорождения ((Bock et al. 1994) или (Garrett 1990)) включают в себя **функциональный** уровень, на котором определяется аргументно-предикатная структура, и **позиционный** уровень, на котором определяется синтаксическая структура. Если в начале 90-ых годов утверждалось, что процессы предикативного согласования всегда контролируются исключительно синтаксическими факторами (т.е. происходят на позиционном уровне), то теперь вопрос о том, могут ли семантические факторы влиять на процессы согласования, чаще решается положительно. Рассмотрим этот вопрос несколько более подробно на примере влияния на процессы предикативного согласования **одушевленности** ГЛИГ и ЛокИГ.

Как и во многих остальных случаях, первый эксперимент, затрагивающий данный вопрос, был представлен в работе (Bock & Miller 1991). В ней авторы анализировали роль категории одушевленности в процессе предикативного согласования: так как одушевленность связана с подлежащностью, то одушевленная ЛокИГ, по их предположению, может сильнее притягивать к себе контроль предикативного согласования, увеличивая при этом вероятность ошибки. Авторы исследовали эту возможность, комбинируя одушевленность и неодушевленность ГЛИГ и ЛокИГ: *The author of the speeches...* (ОдушГЛИГ-НеодушЛокИГ) versus *The speech by the authors...* (НеодушГЛИГ-ОдушЛокИГ). Однако им не удалось обнаружить значительных различий в количестве ошибок в предложениях с данными двумя типами преамбул, что и привело к констатации факта, что в процессе предикативного согласования не происходит обращения к семантике и всё зависит лишь от синтаксиса. Прошло 10 лет и, наконец, в постановке вышеописанного эксперимента была замечена некоторая неточность: почему из четырех возможных комбинаций одушевленности versus неодушевленности и ГЛИГ versus ЛокИГ было проверено лишь две – ОдушГЛИГ-НеодушЛокИГ и НеодушГЛИГ-ОдушЛокИГ? В работе (Barker et al. 2001) авторы выдвинули новую гипотезу: раз предложения с ОдушГЛИГ-НеодушЛокИГ и НеодушГЛИГ-ОдушЛокИГ комбинациями ведут себя одинаково, возможно, что в случае НеодушГЛИГ-НеодушЛокИГ комбинации ошибок будет меньше, чем в случае НеодушГЛИГ-ОдушЛокИГ, и, наоборот, в случае ОдушГЛИГ-ОдушЛокИГ – больше, чем в случае ОдушГЛИГ-НеодушЛокИГ. Проведя подобный эксперимент, авторы обнаружили, что больше всего ошибок возникало в том случае, когда неодушевленная ГЛИГ стоит в Sg, а неодушевленная ЛокИГ – в Pl. Количество ошибок для ОдушГЛИГ-НеодушЛокИГ и НеодушГЛИГ-ОдушЛокИГ, как и в случае эксперимента 1991 года, оказалось примерно одинаковым. Представленные данные свидетельствовали о том, что, хотя первоначальная гипотеза в проведенном исследовании не подтвердилась, процесс предикативного согласования явно чувствителен к категории одушевленности.

### 3. Предикативное согласование в русском языке

В первой работе на материале русского языка (Nicol & Wilson 1999) авторы показали, что падежное маркирование (как еще один дополнительный признак, различающий ГЛИГ и ЛокИГ) существенно уменьшает количество ошибок в процессе предикативного согласования – для русского языка ошибок оказалось существенно меньше, чем для английского. Неожиданным результатом данного эксперимента оказался тот факт, что имена существительные, стоящие в женском роде (Жен), вызывают больше ошибок согласования, чем имена в мужском (Муж): при условии Жен-Жен было зафиксировано 11% ошибок при 2% для остальных условий.

#### 3.1. Эксперимент 1

В эксперименте (Nicol & Wilson 1999) ГЛИГ всегда стояла в Sg; для того, чтобы получить распределение ошибок согласования для всех четырех сочетаний чисел, был смоделирован новый эксперимент на материале русского языка. В нем приняли участие 64 человека, которые заканчивали преамбулы при помощи прилагательного и одного из двух глаголов (*казались* или *оказались*), порождая законченные предложения типа *Компедж рядом с прудами оказался большим*. Стимульный материал эксперимента состоял из 76 преамбул, среди которых было 40 филлеров (отвлекающих предложений), 4 тренировочных и 32 экспериментальных предложения (на каждом из четырех листов был представлен один из четырех вариантов), показанных в (1).

Кроме того, экспериментальный материал был сбалансирован по роду: комбинации Муж-Муж, Муж-Жен, Жен-Муж и Жен-Жен встречались на каждом листе по восемь раз.

Полученные результаты (см. таблицу 1) подтвердили гипотезы о несимметричности распределения ошибок согласования (agreement errors) и о меньшем их количестве по сравнению с английскими данными. Важным результатом эксперимента стало относительно большое количество ошибок повторения (repetition errors), когда испытуемые, неправильно услышав, например, преамбулу *Яма у елок* как *Ямы у елок*, затем правильно ее согласовывали (*Ямы у елок оказались глубокими*).

	ГЛИГ	ЛокИГ	(1)
(большой)	Коттедж рядом с прудом	Sg	Sg
(большой)	Коттедж рядом с прудами	Sg	Pl
(большой)	Коттеджи рядом с прудом	Pl	Sg
(большой)	Коттеджи рядом с прудом	Pl	Pl

Условие ГЛИГ ЛокИГ	Vigliocco_1998_eng		Nicol_1997_eng			rus_experiment_1		
	agr_errors	% rep_errors	agr_errors	%	rep_errors	agr_errors	%	rep_errors
Sg Sg	0,35	1	2,4		1,7	0,4		2,9
Sg Pl	12,5	4,2	13,9		5,9	5,3		8,2
Pl Sg	4,2	9	4,9		2,4	2,5		9,8
Pl Pl	5,2	3,5	4,2		2,4	1,8		9,2
Среднее	5,6	4,4	6,3		3,1	2,5		8

**Таблица 1.** Распределение ошибок согласования и ошибок повторения в зависимости от грамматического числа ГЛИГ и ЛокИГ в экспериментах на английском и русском материале

В результате проведенного статистического анализа подтвердились предварительные результаты, полученные в работе (Nicol & Wilson 1999): при условии Жен-Жен ошибок согласования оказалось значимо больше, чем при других условиях; в двух случаях - Дача через дороги... и Яма у елок... – процент ошибок составил более 30%.

Одним из существенных результатов проведенного эксперимента оказалось выявление того факта, что в случае плохой перцептивной различимости ГЛИГ (как в преамбуле *Дача через дороги...*) возрастает не только количество ошибок повторения, но и количество ошибок согласования. Можно предположить, что именно перцептивная различимость, а не непосредственно родовая характеристика имени, и является причиной преобладания ошибок в случае имен существительных женского рода, которые гораздо чаще в русском языке оказываются трудноразличимыми, чем имена существительные мужского рода. Однако для того, чтобы подтвердить или опровергнуть данную гипотезу, необходимо провести новый эксперимент, в котором будет сбалансировано количество ГЛИГ всех трех родов в зависимости от их перцептивной различимости.

### 3.2. Эксперимент 2

Во втором нашем эксперименте, проведенном по аналогичной методике заканчивания преамбул, приняли участие 72 человека. Стимульный материал эксперимента, сбалансированный по четырем экспериментальным листам, состоял из 100 преамбул, среди которых было 48 филлеров, 4 тренировочных и 48 экспериментальных предложений. Среди экспериментальных предложений было равное количество предложений с ГЛИГ в мужском, женском и среднем (Сред) родах. Все использованные ИГ были неодушевленными, склоняемыми и частотными в современном русском языке. Половина предложений было с МужЛокИГ, половина – с ЖенЛокИГ. Кроме того, для каждого рода ГЛИГ было сбалансировано число перцептивно хорошо различимых и плохо различимых ИГ.

Результаты второго эксперимента еще раз подтвердили гипотезу о несимметричности распределения ошибок согласования – при условии SgPl было допущено 4% ошибок при 0,8% во всех остальных случаях. Как и в предыдущих экспериментах (см. таблицу 2), количество ошибок с ЖенГЛИГ было значимо больше, чем с МужГЛИГ. Главным новым результатом данного эксперимента оказалось количество ошибок с ГЛИГ в среднем роде, которое в несколько раз превысило количество ошибок с двумя другими ГЛИГ. Однако перцептивная гипотеза, выдвинутая по результатам первого эксперимента, во втором эксперименте не подтвердилась – мы не обнаружили значимого различия между количеством ошибок ГЛИГ с плохой и хорошей перцептивной различимостью.

Род ГЛИГ	agr_errors %		
	Nicol 1999_rus	rus_experiment_1	rus_experiment_2
Муж	1	0,01	0,4
Жен	5,6	5	0,9
Сред	-	-	3,6
Среднее	3,3	2,5	1,6

**Таблица 2.** Распределение ошибок согласования в зависимости от грамматического рода ГЛИГ в трех экспериментах, проведенных на русском материале

## 3.3. Обсуждение результатов

Таким образом, и на русском материале мы наблюдаем значительную асимметричность в распределении ошибок при предикативном согласовании – основное их количество приходится на случай, когда ГЛИГ стоит в Sg, а ЛокИГ – в Pl. Однако на русском материале мы обнаружили еще одну важную асимметричность – испытуемые совершают ошибки значимо чаще, когда ГЛИГ имеет форму Жен, чем Муж; в то же время количество ошибок в случае имен среднего рода намного превосходит оба других случая. Как мы можем объяснить полученный эффект?

Получив впервые на русском материале подобное неравномерное распределение ошибок между ЖенГЛИГ и МужГЛИГ, авторы работы (Nicol & Wilson 1999) предположили, что оно может быть вызвано синкретизмом русской падежной системы для имен женского рода. Действительно, анализируя прамбулу *Дача около деревни*, испытуемые могут ошибочно интерпретировать ЛокИГ как стоящую не в родительном падеже Sg, а в именительном падеже Pl, и затем неправильно согласовывать ее - *\*Дача около деревни оказалИСЬ заброшенными*. Подобная гипотеза находит свое подтверждение на материале немецкого языка: в работе (Hartsuiker et al. 2003) показано, что в тех случаях, когда форма ЛокИГ могла восприниматься испытуемыми и как стоящая в именительном падеже и как стоящая в винительном, ошибок наблюдалось значимо больше, чем в случае однозначной ЛокИГ.

Однако на русском материале данная гипотеза выглядит совсем не так убедительно. Во-первых, омонимия появляется только в условии Sg-Sg; между тем, только 5 из 27 ошибок в эксперименте (Nicol & Wilson 1999) совершены при этом условии. Во-вторых, данный синкретизм может действовать и в прамбулах с условием Муж-Жен – *Дом около деревни*. Однако количество ошибок при этом условии совсем невелико. В-третьих, в нашем втором эксперименте только одна ошибка (*Колесо около машины*) может быть потенциально объяснена данной гипотезой, хотя половина ЖенЛокИГ имела подобную неоднозначность. Кроме того, данная гипотеза может объяснить только факт существенных различий между Жен и МужИГ, а нам бы хотелось найти единое объяснение асимметричности всех трех родов в русском языке.

Наше объяснение основано на допущении о том, что сложность порождения высказывания непосредственно зависит от **частотности** тех или иных конкретных лексем или грамматических характеристик, которые используются при его построении. А именно, чем более частотна некоторая грамматическая характеристика, тем легче она обрабатывается носителями языка в процессе порождения высказывания. В случае, когда обработка грамматического рода ГЛИГ вызывает у испытуемого определенные сложности, повышается вероятность интерференции числовой характеристики ЛокИГ на ГЛИГ, тем самым возрастает риск совершения ошибки согласования по числу. Последний вопрос, на который нам остается ответить, состоит в том, какие характеристики ИГ в русском языке более частотны, а какие менее. Традиционно в таких случаях прибегают к помощи языковых корпусов. Из 4,500,000 имен русского национального корпуса ([www.ruscorgora.ru](http://www.ruscorgora.ru)) 49% имен имеют родовую характеристику Муж, 35% - Жен, и только 16% всех имен существительных относятся к Сред. Наименее частотный средний род, согласно нашей гипотезе, должен вызывать наибольшие трудности при согласовании. Таким образом, мы объясняем распределение ошибок предикативного согласования частотностью родовых показателей в языке.

**Список литературы:**

1. Barker, J., Nicol J., & Garrett M. Semantic factors in the production of number agreement. *Journal of psycholinguistic research*, 30, 2001, 91-114.
2. Bates, E. & MacWhinney, B. Functionalism and the competition model. In B: MacWhinney & E. Bates (Eds). *The crosslinguistic study of sentence processing* (pp. 3-73), Cambridge: Cambridge Univ. Press, 1989.
3. Bock, J.K. & Levelt, W.J.M. Language Production. Grammatical Encoding. In M.A. Gernsbacher (Ed.) *Handbook of Psycholinguistics*. San Diego: Academic Press, 1994.
4. Bock, J.K. & Miller, C.A. Broken agreement. *Cognitive Psychology*, 23, 1991, 45–93.
5. Eberhard, K.M. The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, 36, 1997, 147–164.
6. Fodor, J.A. *The modularity of mind*. Cambridge, MA: MIT Press, 1983.
7. Hartsuiker, R.J., Schriefers, H.J., Bock, K., & Kikstra, G.M. Morphological influences on the construction of subject-verb agreement. *Memory and Cognition*, 31, 2003, 1316-1326.
8. Garrett, M.F. Sentence Processing. In D.N. Osherson and H. Lasnik (eds.), *Language: An Invitation to Cognitive Science*, Volume 1. Cambridge MA: MIT Press, 1990.
9. Nicol, J.L., Forster, K.I., & Veres, C. 1997. Subject-verb agreement processes in comprehension. *Journal of Memory and Language*, 36, 569-587.
10. Nicol, J. & Wilson, R. 1999. Agreement and Case-Marking in Russian: A Psycholinguistic Investigation of Agreement Errors in Production. In: FASL, Philadelphia Meeting, 314-327.

11. *Vigliocco, G., Butterworth, B., & Semenza, C.* 1995. Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34, 186–215.
12. *Vigliocco, G., Butterworth, B., & Garrett, M.F.* 1996a. Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61, 261–298.
13. *Vigliocco, G., Hartsuiker, R.J., Jarema, G., & Kolk, H.H.J.* 1996b. One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, 11, 407–442.
14. *Vigliocco, G. & Nicol, J.* 1998. Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, 68, 13–29.