

Социальные сети как источник данных для антропоники¹

Слово и текст в культурном сознании эпохи: Сб. научных трудов. Часть 4. – Вологда: Легия, 2010. С. 97-101.

Отсутствие открытой регулярной статистики по имянаречению в России отмечалось неоднократно (Журавлев 2006: 126). В связи с принятием Закона о персональных данных (О персональных данных 2006) сведения даже общего характера, анонимные по сути, стало невозможным получить ни в государственных органах, ни тем более в коммерческих структурах (Ковалевская 2009).

В то же время бурное развитие социальных сетей в интернете дает новые возможности для изучения русского именика. В социальных сетях (*В контакте* 2006–2010) и (*Одноклассники* 2006–2010) участники, как правило, выступают под своими настоящими именами и указывают достоверные данные о себе. Число пользователей этих сетей достигает десятков миллионов.

Родившиеся в данном городе за определенный год и его современные жители с тем же годом рождения, конечно, суть не одно и то же, особенно если речь идет о крупных городах. Однако русский именик XX века мало меняется от города к городу. Если сравнить верхушку женского именика Пензы 60-х гг. по данным (Бондалетов 1983: 161) с обсуждаемой ниже выборкой загс по Ленинграду 1953–68 гг. (Суслова, Суперанская 1991), то из первых 10 наиболее частотных имен отличие будет в двух именах из нижней части списка. Сейчас, при гораздо более высокой информационной и транспортной связности регионов, отличия эти, на наш взгляд, могли стать только меньше.

В интернет-сети *В контакте* функция поиска человека более развита, чем в сети *Одноклассники*. Она дает возможность задавать ФИО, пол и возраст, регион проживания, учебное заведение, возможен интервальный поиск по возрасту и единичный по дате рождения (отдельно год, месяц и день). При этом выдается и общее число людей, найденных по запросу, которое мы использовали в подсчетах. Проверить валидность полученных таким образом данных и пригодность их для исследований именика и составляет цель нашей работы.

¹ По техническим условиям в материалах конференции статья была опубликована без иллюстраций.

Данные по частотности имен, собранные А.В.Сусловой в ленинградском Дворце малютки (далее называемые *выборками загс*) (Суслова, Суперанская 1991: 83), мы сравнили с цифрами, полученными нами по запросам в сети *В контакте* (далее *сетевые выборки*) в мае 2010 года по городу Санкт-Петербург. Из каждого источника были взяты абсолютные частоты 64 имен девочек, родившихся в 1953–68 и 1988 гг. Размер выборок загс по каждому срезу составляет около 2 тыс носителей имени. Размеры сетевых выборок составили 19 тыс и 29 тыс человек по соответствующим срезам. Корреляцию выборок предполагали линейной.

Коэффициент корреляции Пирсона между выборками из двух источников оказался равным 0,95 для среза 1953–68 гг. и 0,93 для среза 1988 года. Если в сетевой выборке 1988 года учесть имена с их главными гипокористиками (*Наталья+Наташа, Татьяна+Таня*), коэффициент корреляции возрастает до 0,98.

Первая семерка наиболее частотных имен в срезе 1953–68 гг. совпадает на обеих выборках. Первая десятка имен в срезе 1988 года отличается по двум выборкам на одно имя.

Если уравнивать логарифмированием вклады редких и наиболее частотных имен в дисперсию переменных, коэффициент корреляции становится равным 0,87 для среза 1953–68 и 0,93 для среза 1988 без учета гипокористик. Это означает, что относительные частоты имен в рассматриваемых источниках совпадают на 75 и 87% соответственно. Меньшая корреляция на более раннем срезе может быть вызвана неравномерным присутствием разных поколений в сети, рассчитанной на молодую аудиторию: женщин 1953 года рождения зарегистрировано на сайте менее 4 тыс, 1968 года – 91 тыс, 1988 года – 1,2 млн человек. Для более точных исследований корректнее проводить съем данных по узким временным интервалам (3–5 лет) и сравнивать проценты от выборки.

Система календарных имен в России всегда была единой и не связанной с диалектными зонами (Теория и методика... 2007: 60). Единство российского городского именника 60-х гг. XX века показано, в частности, В.А.Никоновым на примере 38 имен по 7 городам (Никонов 1967). По нашим подсчетам, парные коэффициенты корреляции этих массивов лежат в интервале 0,88–0,96 (логарифмическая шкала). Мы сравнили сетевую выборку 1953–68 гг. с данными по этим городам, а также по Москве (Шайкевич 1970: 84) и Свердловску (Короткова 1970: 110). В них всех первая десятка наиболее частотных женских имен отличается от сетевой выборки на 0–2 имени, и всегда это Любовь и Лариса, замещающие 9–10 место, где в сетевой выборке стоят Людмила и Надежда. Единственное исключение – данные по Москве 1953–59 гг., где разница в одно имя и Любовь замещает Светлану сетевой выборки. Первая десятка сетевой

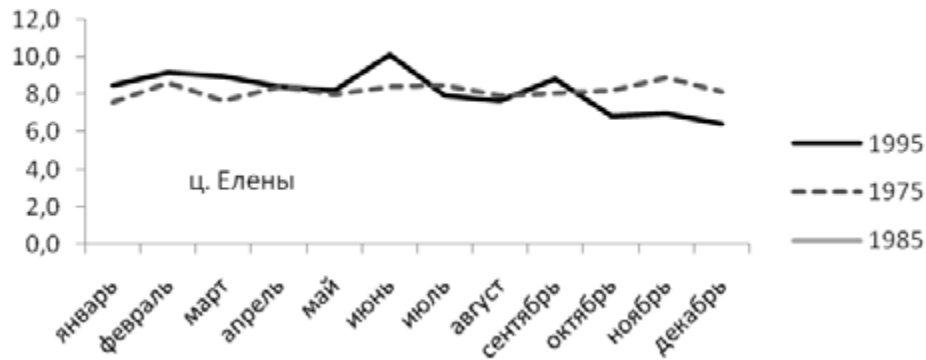
выборки отличается от данных загс (Сулова, Суперанская 1991) также на 2 имени, только здесь это Юлия и Александра (вместо Галины и Надежды в сетевой выборке). Таким образом, первая десятка сетевой выборки демонстрирует большее сходство с данными по другим городам, нежели по Петербургу. Даже без учета первых 10 имен коэффициент корреляции (логарифмическая шкала) сетевой выборки с выборками по городам (кроме Свердловска, где выборка всего 16 имен) не опускается ниже 0,7 и в среднем составляет 0,74.

Следует учитывать, что абсолютные цифры по сети *В контакте* имеют довольно условную природу. Видимо, по запросу с широкими условиями (например, только год рождения и пол) выдаются также и все те, кто скрыл год рождения или не указал пол. Это может исказить картину на редких именах. Но даже в таком брутто-исследовании, как наше, мы получили весьма тесную корреляцию, несмотря на неизбежные искажения данных о себе некоторыми участниками.

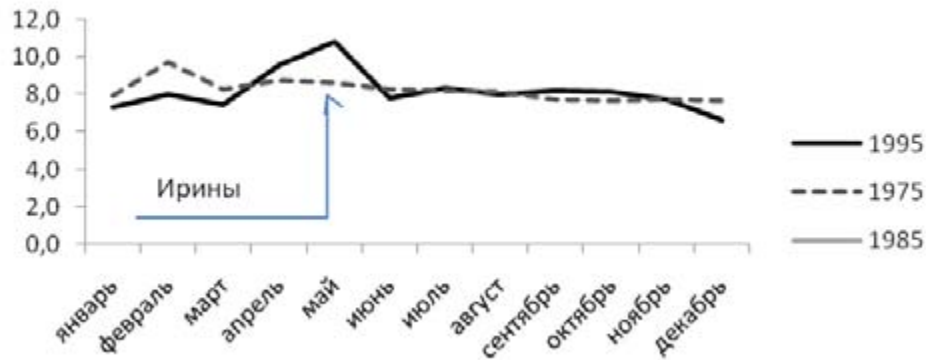
В работах советского периода отмечалось, что в городах изменение частотности имен по месяцам перестало зависеть от церковного календаря (Короткова 1970: 110). Можно предположить возрождение такой связи в современном именнике. В сети *В контакте* мы сняли помесечную динамику имяназвания на нескольких именах по хронологическим срезам 1965–95 гг. с шагом 10 лет. Для имени Татьяна были сделаны дополнительные срезы 1990–94 гг. с шагом в год. Частоты были нормализованы по числу дней в месяце и по числу всех родившихся данном месяце данного года, затем пересчитаны в проценты от выборки данного имени по данному году. Дни памяти святых сверялись по календарю сайта (Православие.ru 1999–2010).

Имя Татьяна показывает в советское время равномерные частоты в течение года на срезах 1965–90 гг., а с 1991 года начинает расти январская частота, достигая в 1995 году 16% против средних 7% по оставшимся месяцам. Для имени Наталья на срезах 1975 и 1995 можно говорить о небольшом повышении частоты в сентябре (день мч. Наталии): 10% против 8% в остальные месяцы; в 1985 году заметного изменения частоты в этом месяце нет. На срезе 1995 года имена Елена, Ирина и Ольга показывают повышенную частотность в месяцы, на которые приходятся дни памяти соответствующих святых.

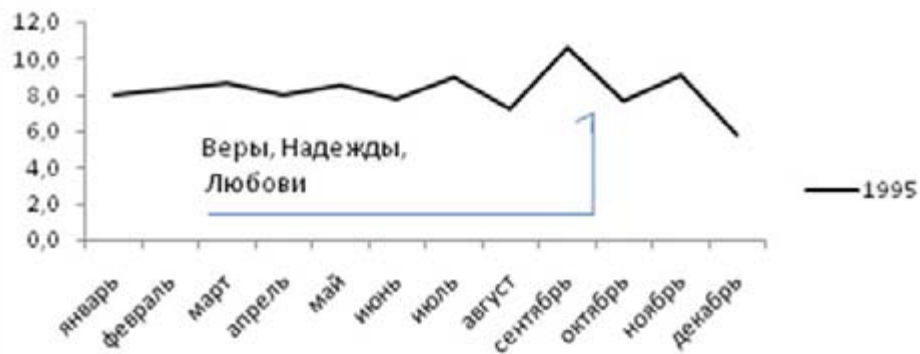
Елена

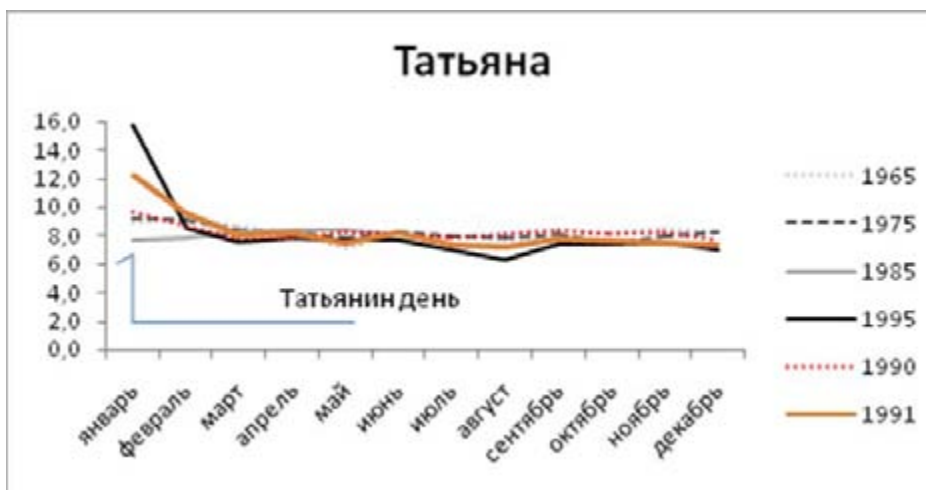
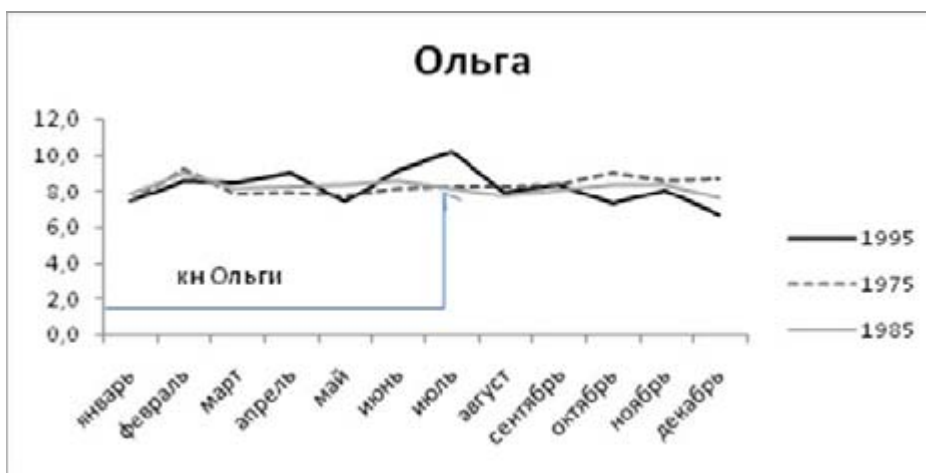
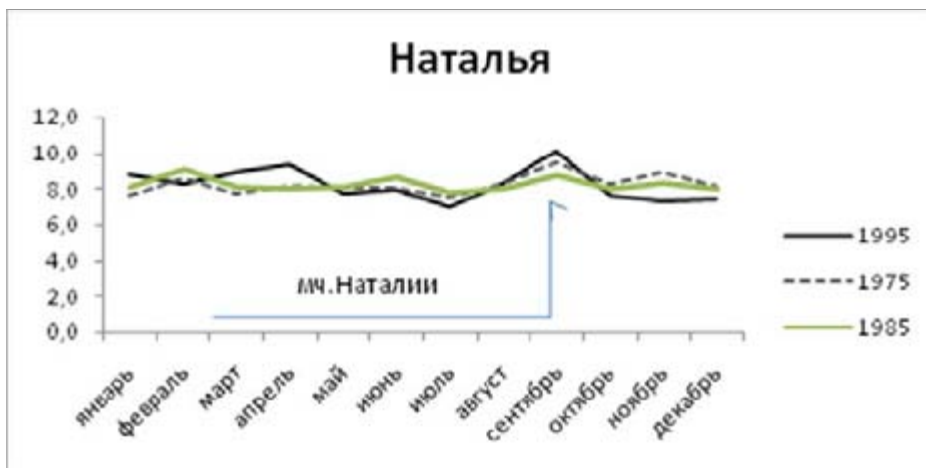


Ирина



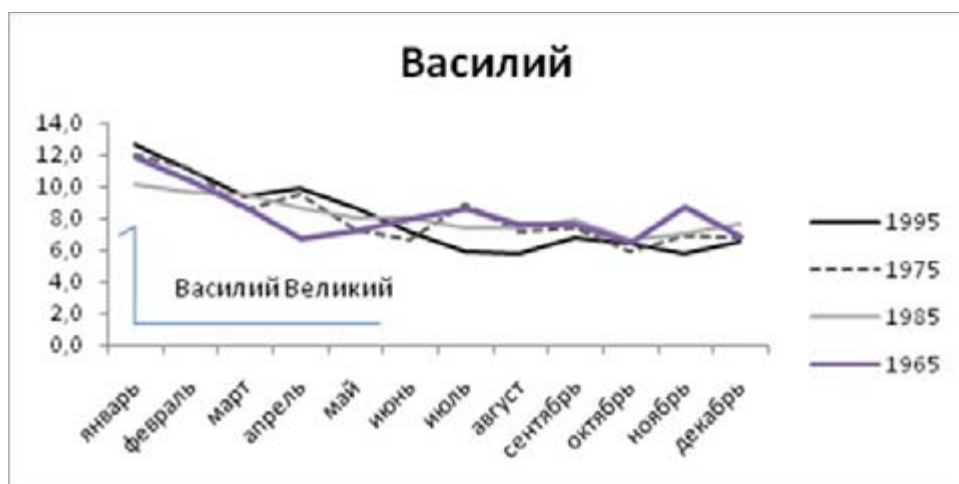
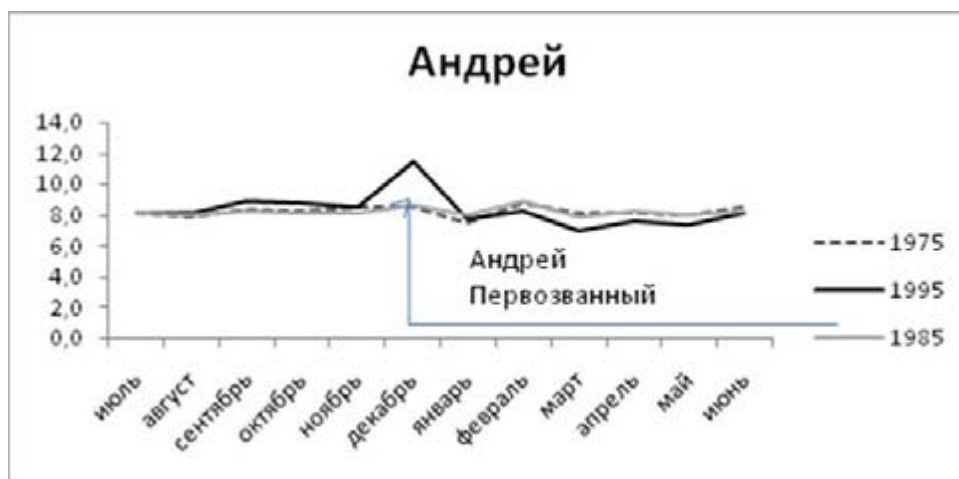
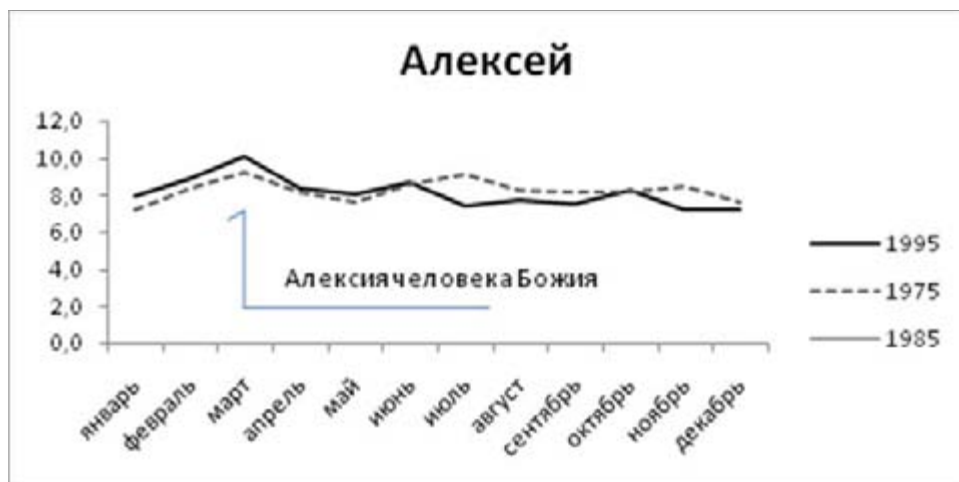
Надежда

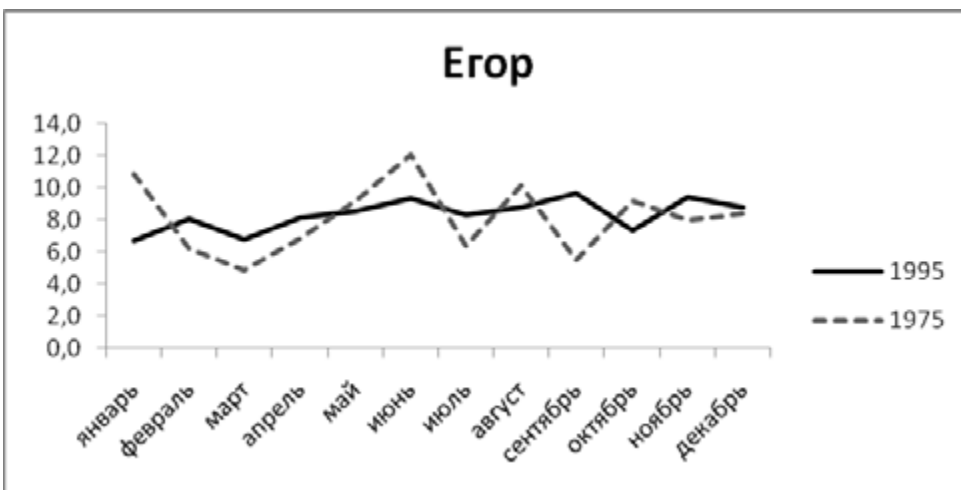
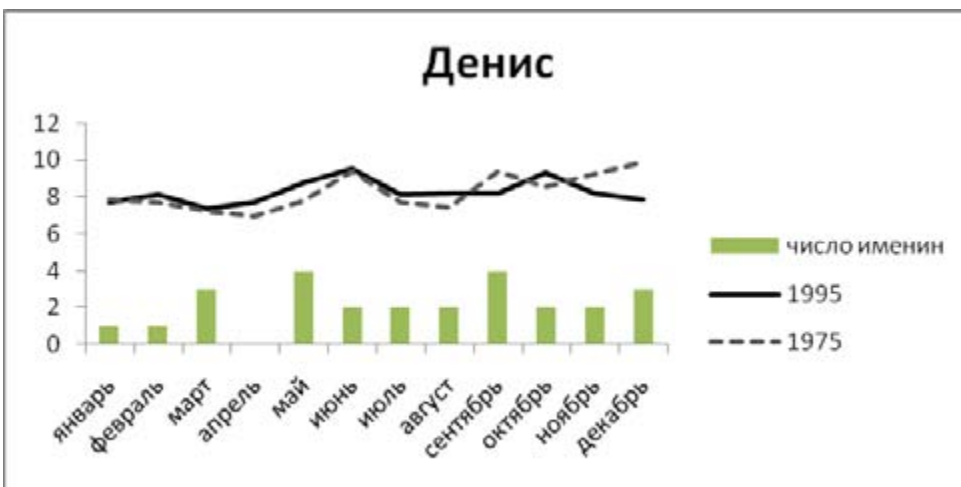
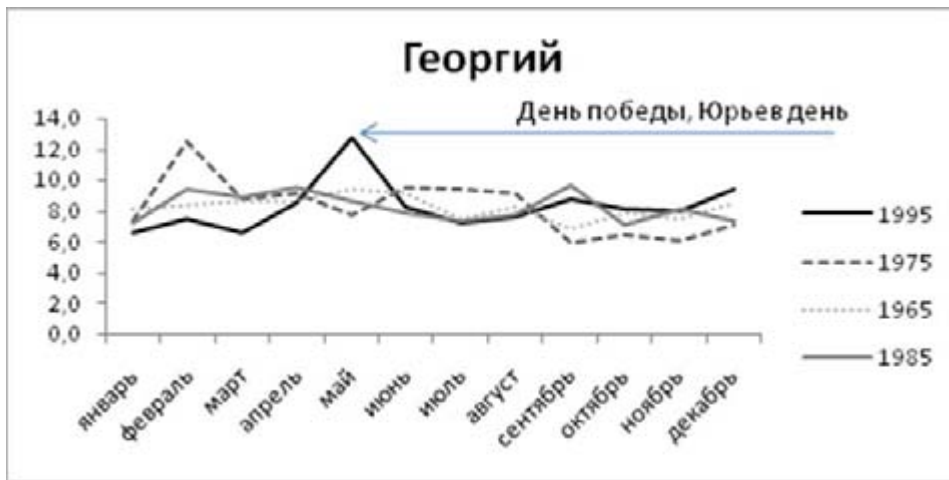




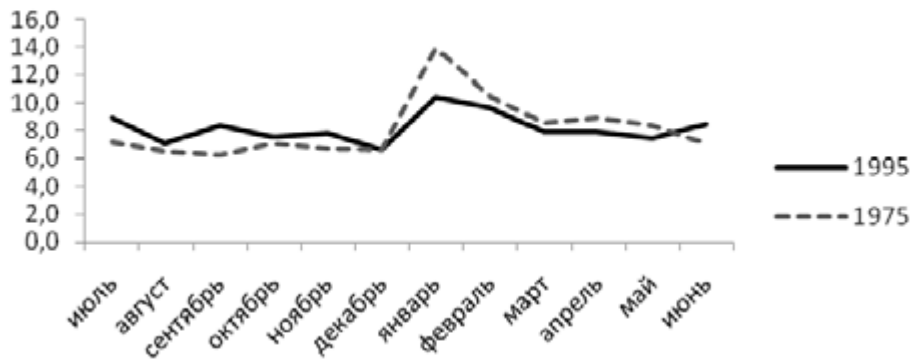
Некоторые мужские имена обнаруживают тесную связь с церковным календарем даже в советское время. Василий в январе, Алексей в марте, Павел (но не Петр) в июле, Илья в августе, Михаил в ноябре на срезях с 1965–95 года показывают заметное повышение частоты. Имя Николай дает два пика частотности, в мае и декабре, что соответствует дням памяти Николая Чудотворца. Имена Денис и Егор дают ровную ежемесячную динамику частоты имянаречения, Андрей и Сергей показывают

повышенные частоты в месяцы, отмеченные днями памяти соименных святых, лишь в 1995 году.



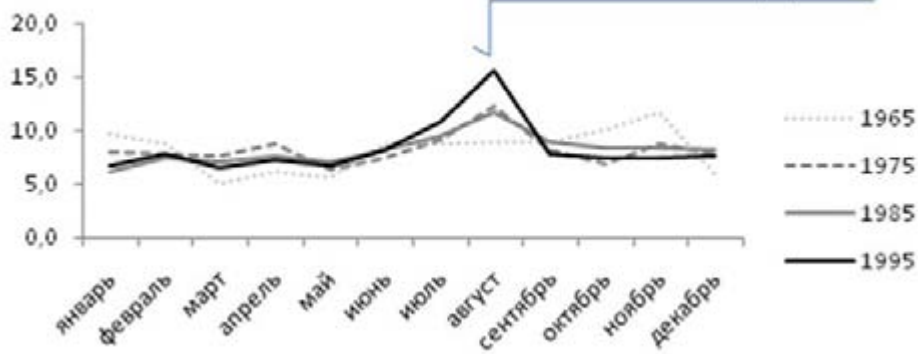


Иван



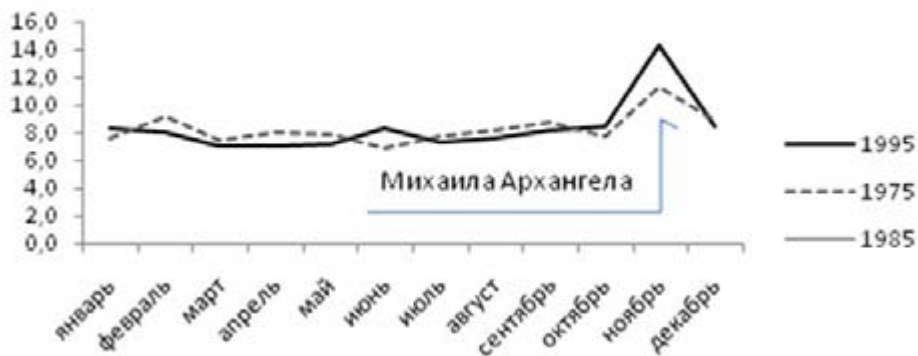
Илья

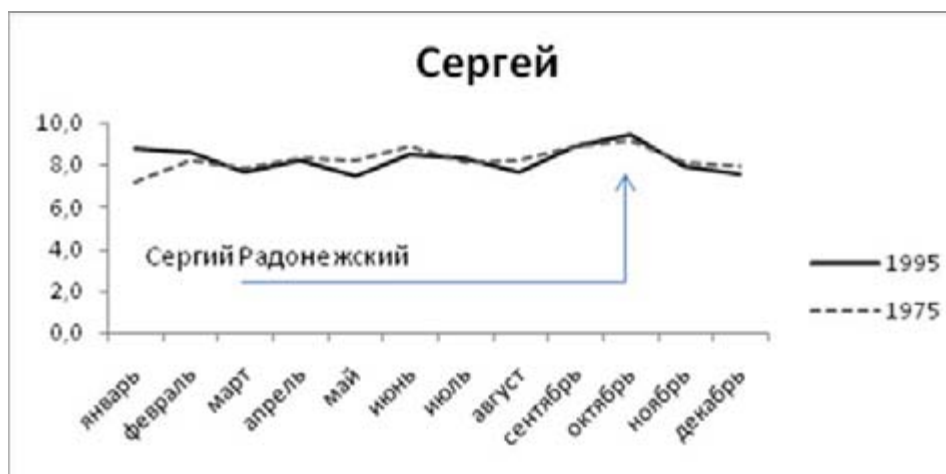
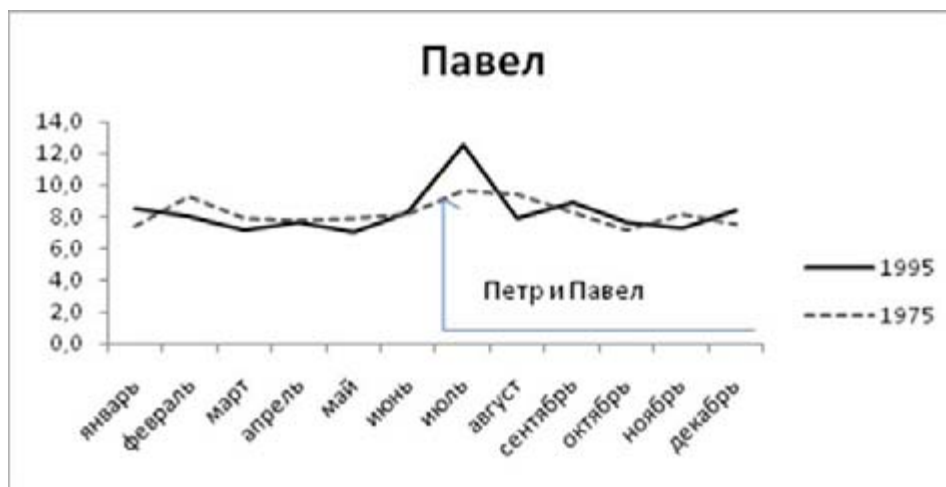
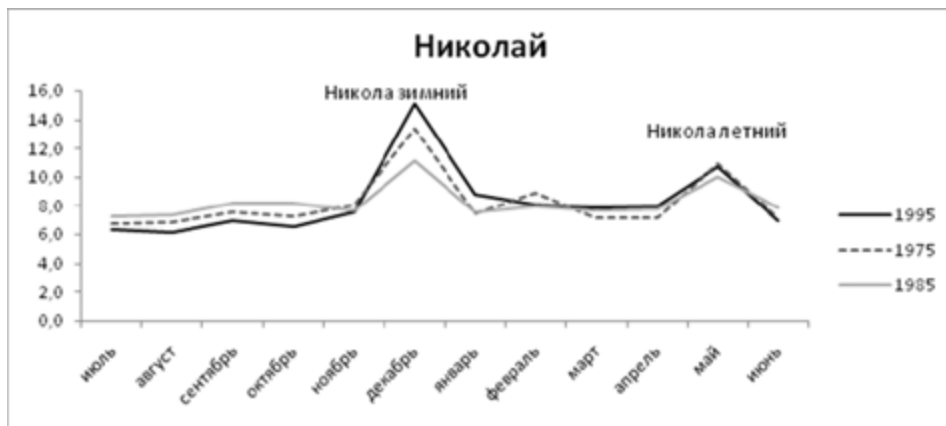
Ильи Пророка

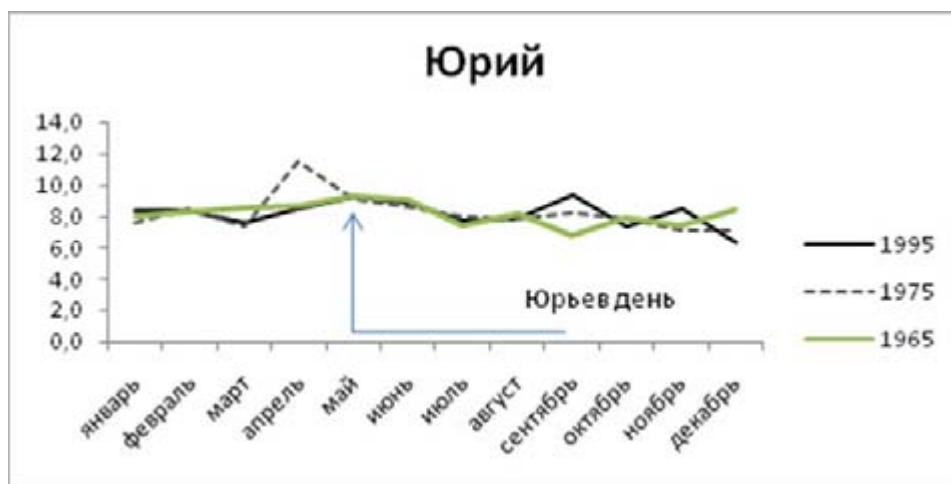


Михаил

Михаила Архангела





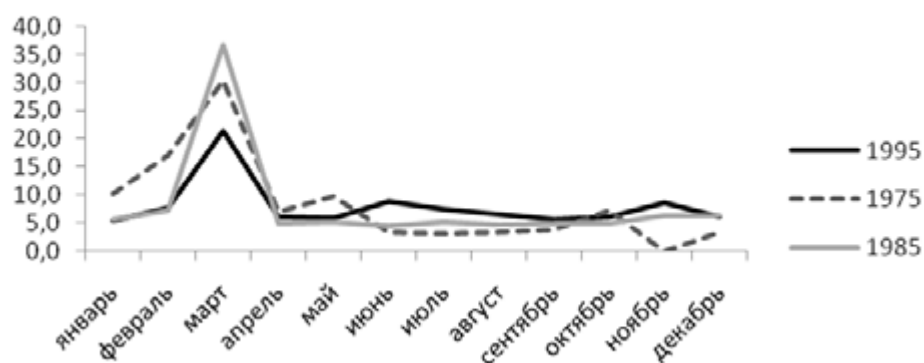


Для сравнения, в Свердловске в 1966 году не отмечено повышения помесечной частотности для имен Ольга, Татьяна, Алексей, Николай (Короткова, 1970). В Брянской области в том же году, наоборот, показана связь с церковным календарем для имен Татьяна, Алексей и Илья (Кондратенко 1970: 115).

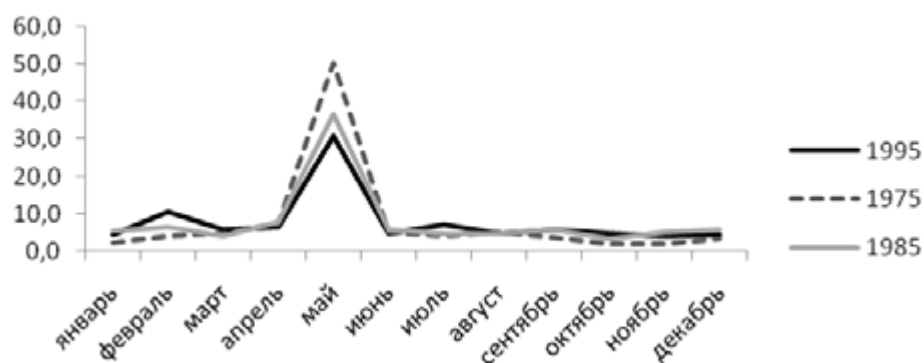
Вообще, на поздних срезах связь имянаречения со святцами кажется несколько сильнее, однако делать выводы на таком небольшом материале преждевременно. Вопрос, почему лишь для части имен наблюдается связь с церковным календарем, независимо от того, насколько часто и равномерно в святцах повторяется имя, заслуживает дополнительного изучения.

Оказалось также, что имена, схожие с названиями месяцев, показывают максимальную частоту рождений в этом месяце. Имя Юлия имеет 10–11% носительниц, родившихся в июле, против среднего по остальным месяцам 8% (срезы 1975–95 гг.). На тех же срезах имя Майя имеет частоту в мае от 31 до 50% (среднее по остальным месяцам 4–6%). Схожая картина выявлена и для имени Марта в марте. Имя Яна показывает повышенную частоту в январе только на срезе 1995 года (15% против среднего 8% по оставшимся месяцам).

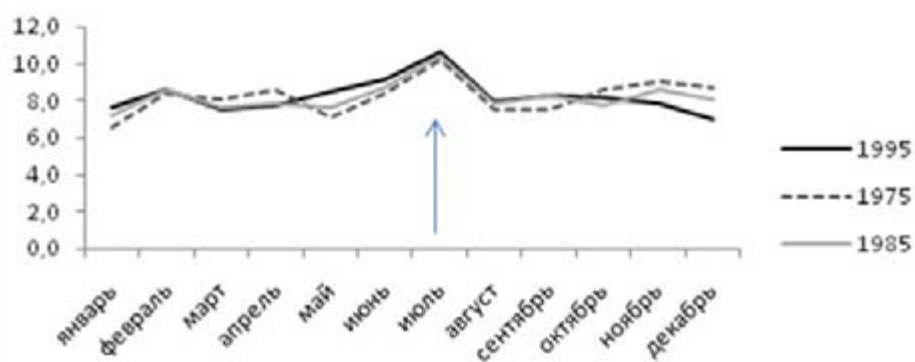
Марта

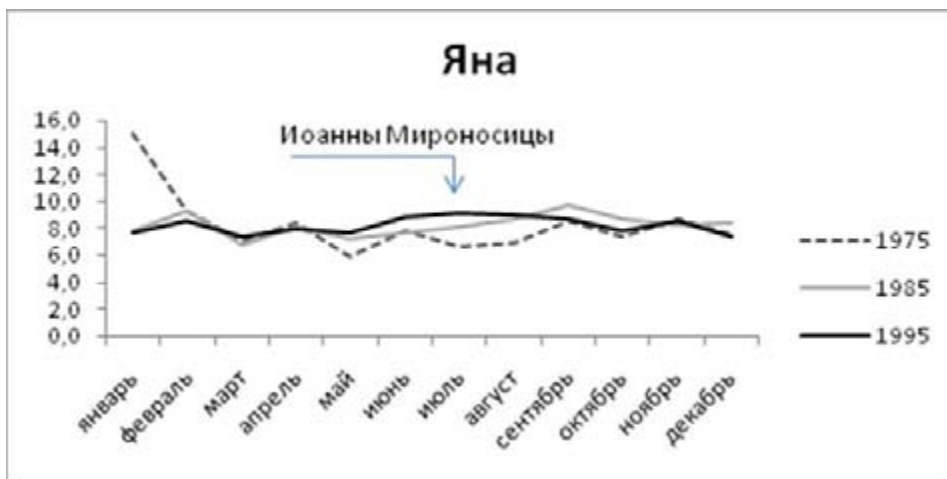


Мая

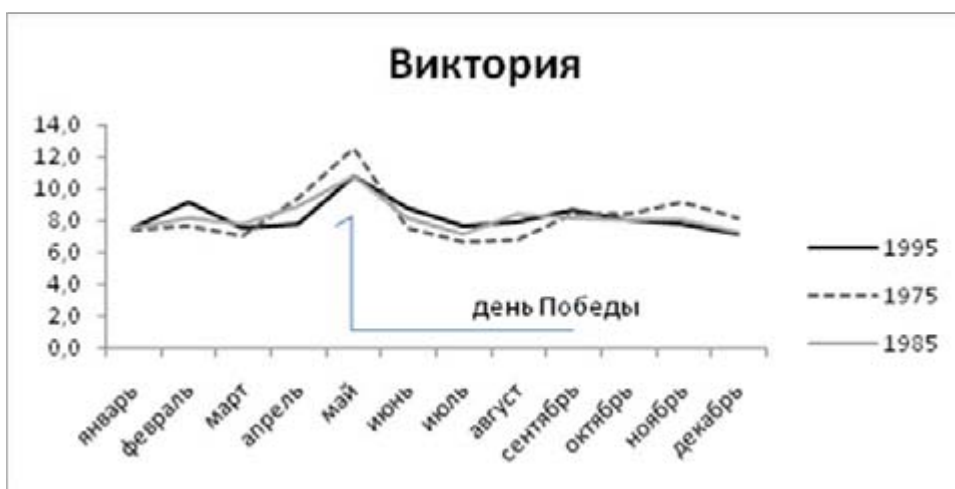
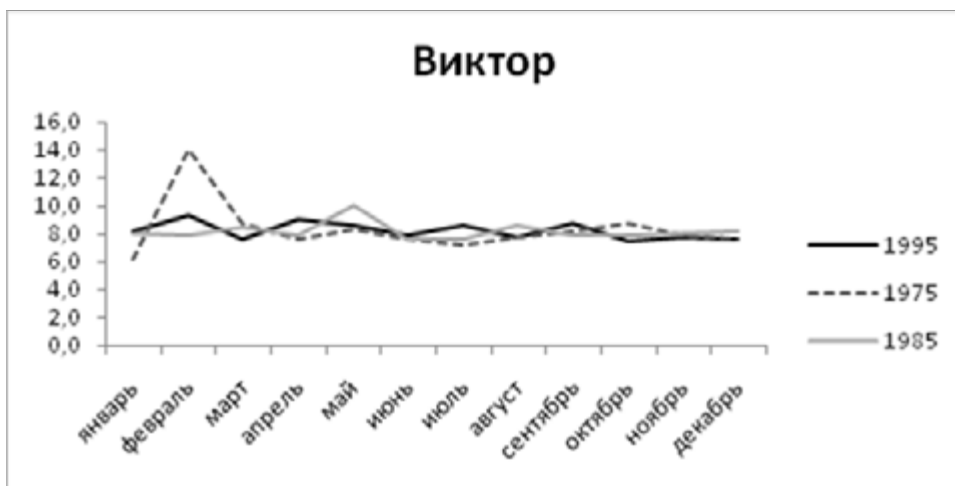


Юлия





Интересно, что имя Виктория дает повышенную частоту в мае (11–12% против средних 8%), что связано, видимо, с днем Победы. Для имени Виктор майского подъема частоты не наблюдается.



Регистрация в сети *В контакте* разрешена лицам, достигшим 14 лет, поэтому для изучения имени младшего поколения следует искать иные

источники. Ими могут стать сетевые сообщества молодых и будущих мам. В них регулярно публикуются собранные самими участницами данные о частотности имен их детей по годам рождения. Проверить репрезентативность этих данных, к сожалению, затруднительно, однако причин для их фальсификации, на наш взгляд, нет. Довольно представительные списки обнаружены на сайтах (sibmama.ru 2001–2010) по Сибири за 1999–2001 гг. и (littleone.ru 2000–2009) по Санкт-Петербургу за нулевые годы. В последнем источнике по нашим 64 именам зарегистрировано 2,4 тыс девочек.

Таким образом, данные по имянаречению из сети *В контакте* обладают достаточной степенью достоверности. Размеры получаемых выборок весьма велики, так что можно проводить статистически значимые измерения. Сказанное верно для крупных городов с высокой степенью проникновения интернета. Вероятно, процесс съема данных несложно автоматизировать и получать таким образом богатый материал для антропонимических исследований.

Источники

Бондалетов В.Д. Русская ономастика. – М.: Просвещение, 1983. – 224 с.

В Контакте. Сетевой проект. – 2006–2010 (Электронный ресурс). URL: www.vkontakte.ru, требуется регистрация (дата обращения 15.05.2010).

Журавлев А.Ф. К статистике русских фамилий. I. // Вопросы ономастики. – 2005. – № 2. – С. 126–146.

Ковалевская Е. Именная грамота. // Рос. газета - Неделя - Кубань-Кавказ. – 2009. – 30 апреля. URL: www.rg.ru/2009/04/30/reg-kuban/imena.html (дата обращения: 25.05.2010).

Кондратенко Г.И. Наблюдения над употреблением личных имен в Брянской области. // Личные имена в прошлом, настоящем, будущем: проблемы антропонимики. / Ответ. ред. В.А.Никонов. – М.: Наука, 1970. – 344 с.

Короткова Т.А. Динамика личных имен свердловчан. // Личные имена в прошлом, настоящем, будущем: проблемы антропонимики. / Ответ. ред. В.А.Никонов. – М.: Наука, 1970. – 344 с.

Никонов В.А. Личные имена в современной России. // Вопросы языкознания. – 1967. – №6. – С. 102–111. Цит. по: Русское личное имя /

Википедия: свободная энциклопедия. – 2010 (Электронный ресурс). URL: ru.wikipedia.org/wiki/Русское_личное_имя (дата обращения: 29.05.2010).

О персональных данных: федер. закон Рос. Федерации от 27 июля 2006 г. N 152-ФЗ. // Рос. газета. – 2006. – 29 июля. URL: www.rg.ru/2006/07/29/personaljnye-dannye-dok.html (дата обращения: 20.05.2010).

Одноклассники. Сайт. – 2006–2010 (Электронный ресурс). URL: www.odnoklassniki.ru. Требуется регистрация (дата обращения 29.05.2010).

Православие.ru: по благосл. Святейшего Патриарха Моск. и всея Руси Алексия II. – 1999–2010 (Электронный ресурс). URL: days.pravoslavie.ru/ABC/wn.htm (дата обращения 29.05.2010).

Суслова А.В., Суперанская А.В. О русских именах. – Л.: Лениздат, 1991. – 220 с.

Теория и методика ономастических исследований. / Суперанская А.В., Сталтмане В.Э., Подольская Н.В., Султанов А.Х. Отв. ред. А.П. Непокупный. – М.: Издательство ЛКИ, 2007. – 256 с.

Шайкевич А.Я. Русские личные имена XX в. (по материалам ЗАГСа Фрунзенского района г. Москвы). // Личные имена в прошлом, настоящем, будущем: проблемы антропонимики. / Ответ. ред. В.А.Никонов. – М.: Наука, 1970. – 344 с.

littleone.ru: Клуб питерских родителей. – 2000–2009 (Электронный ресурс). URL: www.littleone.ru/club/names/name_test.php (дата обращения 29.05.2010).

sibmama.ru: Сибирский семейный сайт. – 2001–2010 (Электронный ресурс). URL: sibmama.ru/index.php?p=names_2001 (дата обращения 29.05.2010).

* * *