

Крючкова О.Ю. ([vpks@rambler.ru](mailto:vpks@rambler.ru))

Саратовский государственный университет  
имени Н.Г. Чернышевского

**ВКЛАД**

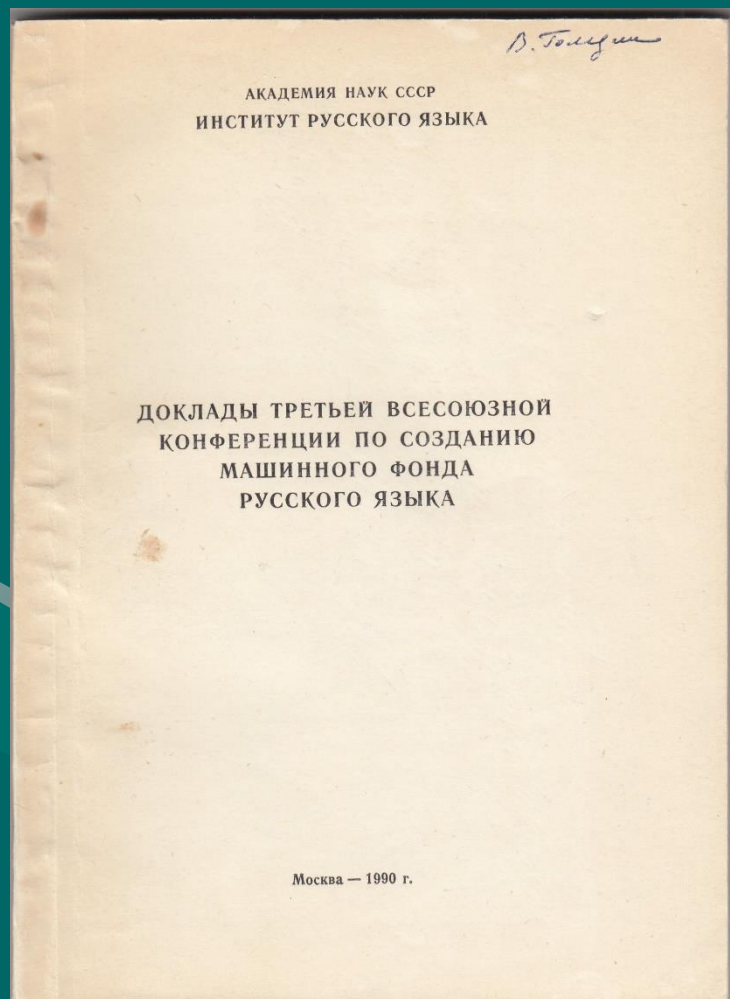
**ВАЛЕНТИНА ЕВСЕЕВИЧА ГОЛЬДИНА**

**В РАЗВИТИЕ КОРПУСНОЙ  
ДИАЛЕКТОЛОГИИ**

# Валентин Евсеевич Гольдин (1935 – 2017)



# III Всесоюзная конференция по созданию машинного фонда русского языка (М., 1989 г.)



А.С. Гердом и В.Е. Гольдиным были высказаны идеи о создании диалектологического подфонда МФРЯ:

- об ориентации данного подфонда на полные магнитофонные записи диалектных текстов (А.С. Герд, В.Е. Гольдин);
- о текстовых корпусах отдельных говоров как основе диалектологического подфонда МФРЯ (В.Е. Гольдин).

## В.Е. Гольдин о значении диалектологического подфонда в составе МФРЯ

«Диалекты занимали в прошлом и занимают в составе современного русского языка такое важное место, что возможности использования МФ русского языка в фундаментальных исследованиях и прикладных целях в значительной мере зависят от того, как представлен в нем материал русских народных говоров»

(Гольдин В. Е. К проекту текстового диалектологического подфонда Машинного фонда русского языка // Доклады III Всесоюзной конференции по созданию машинного фонда русского языка. М.: Ин-т русского языка АН СССР, 1990)

## В.Е. Гольдин о двух типах диалектологических баз данных в составе МФРЯ

- Материал русских народных говоров должен войти в МФРЯ как в виде диалектологических атласов, словарей, картотек (вторичные источники – речевой материал, организованный в соответствии с определенными научными задачами), так и виде речевых произведений на диалекте – текстов (первичные источники).
- Представление двух типов источников в диалектологическом подфонде НКРЯ будет коррелировать с представлением в МФРЯ фактов русского литературного языка.

# Почему корпус текстов «должен составлять основу диалектологического подфонда МФ русского языка»?

## Вторичные источники для литературного языка

### vs. вторичные источники для диалекта

- ✓ Вторичные по отношению к литературному языку источники «создаются и действуют внутри литературного общения, внутренне ему присущи, органичны для него»;
- ✓ «вторичные источники изучения диалектной речи имеют по отношению к диалектному общению абсолютно внешний характер: эти тексты не обращаются в системе диалекта <...>, они создаются лицами, обычно не владеющими диалектом практически, создаются не на диалекте, не становятся известными диалектоносителям (усвоение диалекта происходит не с их помощью) и почти не оказывают воздействия на диалект...» [Гольдин 1990]



Конечная цель разработки диалектной текстовой БД –  
создание текстового фонда  
«обобщающей направленности»

Текстовый диалектный фонд «обобщающей направленности» должен отражать «важнейшие явления русской народной речи в соответствии с единицами диалектного членения языка (наречиями, зонами, группами...) по территориям исконного и позднего заселения» [Гольдин 1990].

## В.Е. Гольдин о двух типах диалектологических корпусов

1. Корпус текстов русской диалектной речи («должен служить машиннообрабатываемой научной хрестоматией русских народных говоров»).

Структура 1-го типа корпуса обусловлена характером диалектного членения русского языка, корпус должен представлять единицы данного членения.

2. Обширные текстовые корпуса некоторых отдельных говоров.

Выбор конкретного говора для ее корпусного воплощения «может определяться рядом «внешних обстоятельств»: объемом и качеством накопленного материала, отражением в материале различных активно взаимодействующих групп населения, отражением присущих диалектному общению на данной территории речевых событий, жанров, актов.

*(Гольдин В. Е. Диалектологический текстовый машинный фонд говора и исследование диалектных изменений // Современные процессы в русских народных говорах. Саратов: Изд-во Саратов. ун-та, 1991)*



# Теория коммуникативной диалектологии – основа корпусной концепции В.Е. Гольдина

Гольдин Валентин Евсеевич

**Теоретические проблемы  
коммуникативной диалектологии**

Специальность 10.02.01 – Русский язык

Диссертация в виде научного доклада,  
представленная на соискание ученой степени  
доктора филологических наук

Саратов  
1997

- Русские народные говоры целесообразно изучать не только в лингвогеографическом аспекте, но и как самостоятельные коммуникативные системы, представляющие особый тип речевой культуры;
- «целесообразно изучать не только междиалектные различия, но и признаки, общие всем диалектам, – признаки диалекта как особого типа речи».

# Новые объекты коммуникативной диалектологии

- информационная структура общения на диалекте (характерный для диалектной речи набор речевых событий, речевых жанров и др.);
- особенности текстовой деятельности на диалекте (особенности построения текстов);
- когнитивная сторона общения на диалекте (особенности картины мира носителей традиционной деревенской культуры);
- место речи в составе деятельности носителей традиционной культуры и характер рефлексии диалектоносителей над речью;
- особенности сохранения речевой традиции, специфика трансляции ее во времени, характер прецедентных текстов диалектного общения.

# Тексты – основной источник коммуникативной диалектологии



Решение задач коммуникативной диалектологии требует специальной организации научных источников: главным объектом наблюдений становятся связная речь и воплощенные в ней единицы общения, тексты.



## Дисбаланс в изучении литературной и диалектной речи определяется во многом «формой фиксации и представления научных фактов»

Если «русская литературно-книжная речь традиционно изучается на материале фиксированных текстов», то источниками диалектологии чаще всего становятся «картотеки, списки, словники и тому подобные вторичные базы данных, элиминирующие текстовые (и тем более – дискурсные) связи в материале»

*(Гольдин В.Е. Машиннообрабатываемые корпуса диалектных текстов и проблема типологии русской речи // Русистика сегодня. 1995. №3)*

# Принципы построения диалектного корпуса

- Корпус представляет реально функционирующие коммуникативные образования – говоры как диалектные микросистемы.
- Корпус представляет диалект как целостное культурно- коммуникативное образование.

## Принципы построения диалектного корпуса

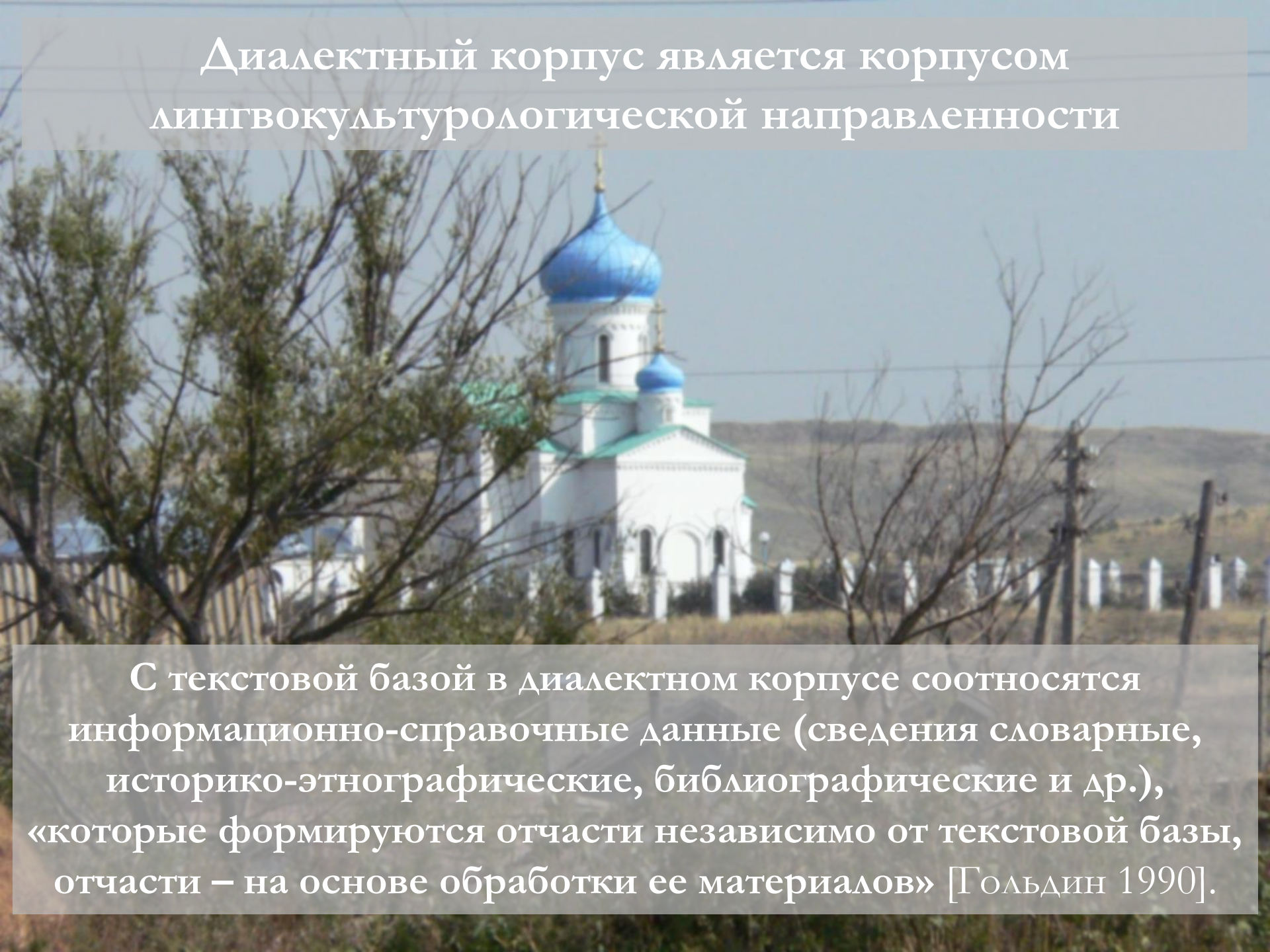
### Текстовая база диалектного корпуса отражает:

- важнейшие типы диалектной речи
- различные формы речи
- разнообразную тематику сельского общения
- социальную дифференциацию носителей говора
- типы говорящих (языковые личности носителей диалекта)
- типы прецедентных текстов





# Диалектный корпус является корпусом лингвокультурологической направленности



С текстовой базой в диалектном корпусе соотносятся информационно-справочные данные (сведения словарные, историко-этнографические, библиографические и др.), «которые формируются отчасти независимо от текстовой базы, отчасти – на основе обработки ее материалов» [Гольдин 1990].

# О формах представления текстов в диалектном корпусе (развитие подходов)

- **1990 г.:** Тексты представляются в корпусе «в подробной фонетической транскрипции, в ВТ (во вспомогательной транскрипции – О.К.) и в аналоговой форме» [Гольдин 1990].
- **2008 г.:** Основная единица корпуса – «запись»/ «текст» – хранится в трех видах:
  - в виде звукового модуля,
  - в виде текстового модуля с символьной расшифровкой аудиозаписи,
  - в виде текстового модуля с размеченным текстом.

(Крючкова О.Ю., Гольдин В.Е. Текстовый диалектологический корпус как модель традиционной сельской коммуникации // Компьютерная лингвистика и интеллектуальные технологии. Вып. 7 (14). М.: РГГУ, 2008)

## Фрагмент размеченного диалектного текста (морфологическая, тематическая, жанровая разметка)

#9@1\*2 сватья {сватья(сватья)=S,жен,од=ед,им} у {у(у)=PR}  
нас {мы(мы)=S-PRO,мн,од=род} тоже {тоже(тоже)=ADV}  
скончалася {скончаться(скончаться)=V=сов,изъяв,прош,ед,муж=nstand}  
{дак(дак)=PART=nstand} /  
двоюродна {двоюродный(двоюродный)=A=ед,жен,им,стяж=nstand}  
сестра {сестра(сестра)=S,жен,од=ед,им} моя {мой(мой)=A=ед,жен,им}  
тоже {тоже(тоже)=ADV} / причитывала {причитывать(исполнять  
песни-плачи)=V=несов,изъяв,прош,ед,жен=nstand} / а {а(а)=CONJ}  
ей {она(она)=S-PRO,жен,од,ед=дат} вот {вот(вот)=PART}  
Наташа {Наташа(Наташа)=S,name,жен,од=ед,им} у {у(у)=PR}  
нас {мы(мы)=S-PRO,мн,од=род}  
уме {уметь(уметь)=V=несов,изъяв,наст,ед,3-л=nstand}  
причитывать {причитывать(исполнять песни-  
плачи)=V=несов,инф=nstand} // %A как же этому учатся?%  
ну {ну(ну)=PART} / не {не(не)=PART}  
учатся {учиться(учиться)=V=несов,изъяв,наст,мн,3-л}  
да {да(да)=PART} ... %Или просто слушают?%  
прислухивают {прислухивать(запоминать на  
слух)=V=несов,изъяв,наст,мн,3-л=nstand} // @1\*2#9

(Волог. обл., Мегра)



## В.Е. Гольдин об эвристическом потенциале текстовых диалектных корпусов

Новые возможности корпусного исследования говорков определяются:

- «значительной полнотой сохранения в источнике (базе данных) реальных отношений между речевыми единицами потока диалектной речи»;
- «преимуществами автоматизированного управления базой данных (получением разного рода сортировок, прямых и обратных, частотных словарей, конкордансов и т.п.)»;
- долгосрочной перспективой расширять предметное поле исследования: текстовая форма фиксации и хранения диалектного материала позволяет исследователю анализировать любые представленные в тексте единицы «независимо от того, планировалось ли обращение к ним на этапе сбора материала, осознавалась ли их важность в ходе полевой работы»

(Гольдин В. Е. Диалектологический текстовый машинный фонд говора и исследование диалектных изменений // Современные процессы в русских народных говорах. Саратов: Изд-во Саратов. ун-та, 1991)



Концепция корпусной диалектологии, выдвинутая В.Е. Гольдиным в конце 1980-х — нач. 1990-х гг., получила дальнейшее развитие при создании диалектологического корпуса в Саратовском государственном университете - СарДК.

## Некоторые работы по СарДК:

- *Крючкова О.Ю., Гольдин В.Е.* Текстовый диалектологический корпус как модель традиционной сельской коммуникации // Компьютерная лингвистика и интеллектуальные технологии. Вып. 7 (14). М., 2008.
- *Крючкова О.Ю., Гольдин В.Е.* Корпус русской диалектной речи: концепция и параметры оценки // Компьютерная лингвистика и интеллектуальные технологии. Вып. 10 (17). М., 2011.
- *Крючкова О.Ю., Гольдин В.Е.* Параметры обработки текстов для русского диалектного корпуса // Труды международной конференции «Корпусная лингвистика— 2015». СПб., 2015.