

Системы машинного перевода.

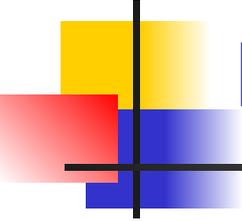
Борисов В.А.

КАСК – филиал ФГБОУ ВПО РАНХ и ГС
Красноармейск 2011 г.



***СРЕДСТВА
АВТОМАТИЗАЦИИ
ПЕРЕВОДОВ***

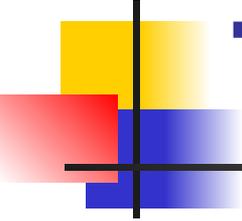
Средства автоматизации перевода текстов



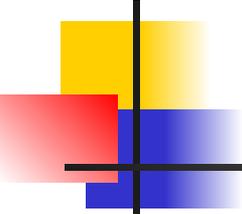
- текстовый редактор;
- электронные словари;
- системы Translation Memory (TM);
- технологии машинного перевода (МП).

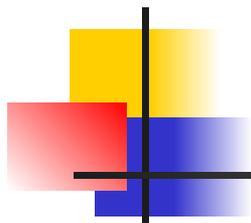
Принцип работы систем

ТМ



- Построен на сравнении текста, подлежащего переводу, с внутренней базой, хранящей разбитый на предложения параллельный текст на исходном языке и его перевод, полученный традиционным способом.

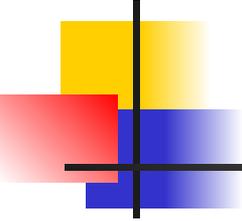
- 
-
- Наиболее известными системами ТМ в настоящее время являются Trades TWB и Transit.
 - Наибольшую эффективность обеспечивает комплексное решение, опирающееся на все имеющиеся в настоящий момент лингвистические технологии машинного перевода.



ИСТОРИЯ ЭЛЕКТРОННОГО ПЕРЕВОДА

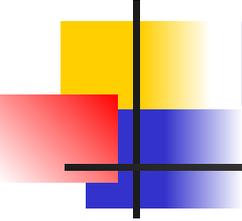
- 
-
- Работы по машинному переводу начались с середины 1950-х годов.
 - Наибольшую известность приобрели труды японского ученого М. Нагао, который предложил использовать при машинном переводе уже готовые, введенные в память ЭВМ варианты прочтения исходных текстов, ранее сделанные опытными лингвистами.

- 
-
- Активные исследования в области машинного перевода велись и за рубежом силами таких известных компьютерных лингвистов, как А. Эттингер (США), И.Бар-Хиллер (США), Н. Хомски (Израиль).

- 
-
- У нас в стране пионерами теории машинного перевода были Д. Ю. Панов, А. А. Ляпунов и И. С. Мухин.

ОТЕЧЕСТВЕННЫЕ СИСТЕМЫ МАШИННОГО ПЕРЕВОДА

- Российскими разработчиками созданы четыре поколения программ-переводчиков.



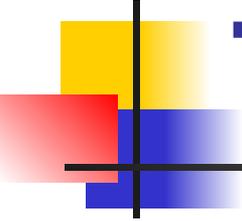
Первое поколение

- 1991 г. - появление пакета PROMT;
- 1992 г. - появился пакет Stylus;
- 1993 г. - был создан пакет Stylus for Windows 2.0.



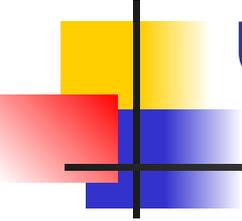
Второе поколение

- В 1994 г. была представлена полноценная версия Stylus for Windows;
- В 1995 г. был разработан пакет Stylus Lingvo Office.



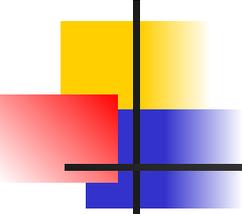
Третье поколение

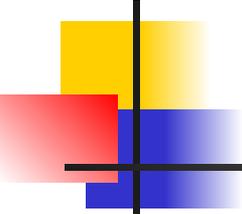
- В 1996 г. появляется 32-разрядная версия Stylus под Windows-95 и –NT;
- Переводчики для Internet — WebTranSite и «упрощенная» версия Stylus Lite появляются в 1997 г.

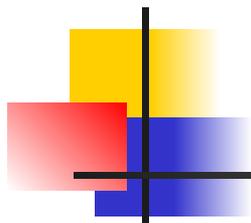


Четвертое поколение

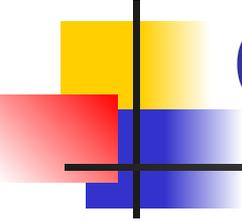
- Программа-переводчик четвертого поколения PROMT-98:
 - PROMT - среда переводчика;
 - File Translator - приложение для пакетной обработки большого количества документов;
 - WebWiew — браузер с синхронным переводом HTML-страниц.

- 
-
- В основе программ-переводчиков четвертого поколения лежит технология HTML-to-HTML, позволяющая переводить Web- страницы с полным сохранением форматирования и впоследствии двигаться по переведенным ссылкам.

- 
-
- В 1998 г. появляется новая версия переводчика для Интернета — WebTranSite-98 и версия 1.0 переводчика для карманных компьютеров — Pocket PROMT.



***ПЕРЕВОДЧЕСКИЕ
ПАКЕТЫ
PROMT***



Состав пакета PROMT

- SmarTool,
- Mail Translator,
- PROMT,
- Dictionary Editor,
- Electronic Dictionary,
- WebWiew,
- QTrans,
- Clipboard Translator и Integrator.

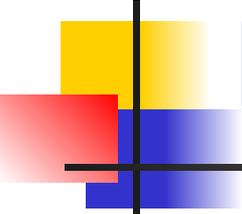


Magic Goody

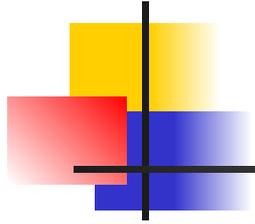
- Кроме обычных для переводчиков возможностей умеет выполнять команды, которые вы произнесете вслух (на русском и английском языках), произносить текст или его перевод на английском или русском языках, переводить текст практически из любого приложения или редактора.

Программы-переводчики

PROMT

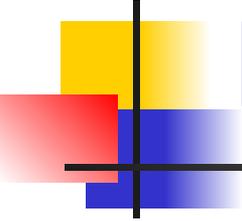


- предоставляют широкие возможности по настройке на перевод текстов, подключению специализированных словарей, созданию пользовательских словарей;
- текст для перевода может вводиться не только с клавиатуры, но и со сканера, а также возможен перевод текстов непосредственно в текстовых редакторах.



Особенности работы программы PROMT

Главное окно программы PROMT

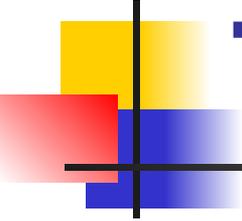


- Разделено на две части, под которыми располагается информационная панель.
- Для перевода текста исходный текст помещается, набирается или копируется в верхнюю часть экрана, нижняя часть предназначена для переведенного текста.



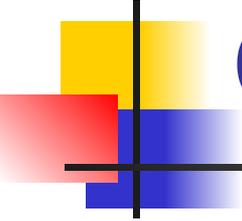
Информационная панель

- На вкладке *Используемые словари* отображается список подключенных словарей, которые будут использоваться при переводе.
- На вкладку *Незнакомые слова* будут помещаться слова, которые не найдены в словарях при переводе.
- Некоторые слова не требуют перевода, и их надо зарезервировать путем перетаскивания на третью вкладку информационной панели - *Зарезервированные слова*.



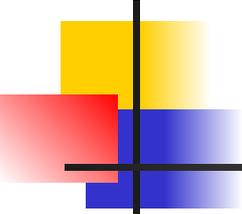
Транслитерация

- Запись слова одного языка буквами другого, например Internet — Интернет.



Словари PROMT

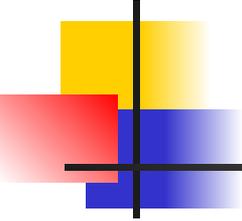
- генеральные,
- специализированные,
- пользовательские.



- *Специализированные словари*

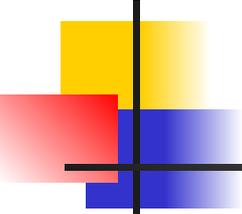
предназначены для перевода текста по определенной тематике, а

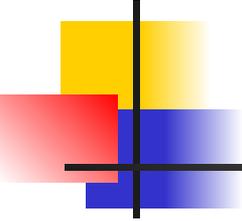
пользовательские словари создаются самими пользователями исходя из их опыта перевода.

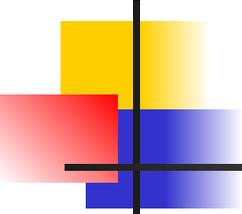
- 
-
- При переводе система PROMT использует только те словари, которые помещены в список «Используемые словари», отображаемый на информационной панели.

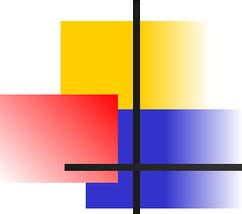


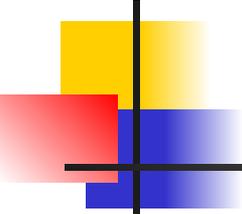
***Последовательность
действий
при выполнении
перевода в PROMT***

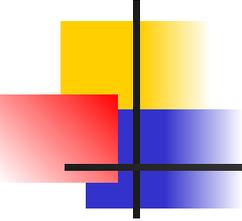
- 
-
- Открыть файл с исходным текстом или создать новый документ, если текст будет набран непосредственно в PROMT.
 - Исходный текст появится в окне документа.

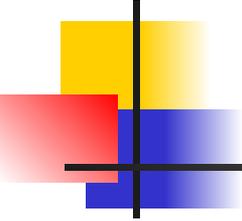
- 
-
- Отредактировать исходный текст, если это необходимо, обратив внимание на разбивку текста на абзацы.

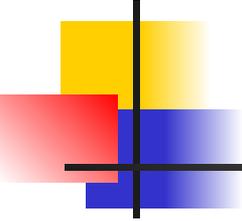
- 
-
- Подключить словари, которые будут использоваться при переводе текста.
 - Это позволит получить более качественный перевод предметно-ориентированных текстов.

- 
-
- Для изменения списка подключенных словарей щелкните правой кнопкой мыши по вкладке Используемые словари и выберите команду Изменить список словарей либо выберите в меню Словари команду, указывающую языковую ориентацию словарей, например Англо - Русские.

- 
-
- В случае необходимости произведите резервирование определенных слов, чтобы некоторые слова не переводились и оставались в тексте перевода так, как они встречаются в исходном тексте.
 - В программе предусмотрена возможность отметки целых абзацев, которые не требуют перевода.

- 
-
- Осуществить непосредственный перевод текста.
 - Все команды, выполняющие перевод текста, находятся в меню Перевод.

- 
-
- Можно переводить текст по абзацам или сразу весь текст.

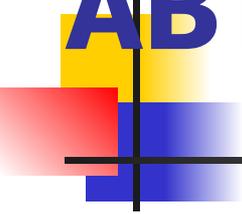
- 
-
- Можно пользоваться мгновенным переводом слова: система автоматически высвечивает перевод слова, на котором установлен курсор, при этом никакой команды вызывать не надо.

Сохранение

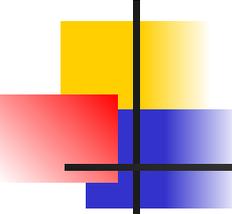
переведенного текста

- Для сохранения полученного текста перевода выберите команду Сохранить/Перевод из меню Файл.
- В списке Тип выберите желаемый формат для сохранения результатов перевода: текстовый или RTF, укажите папку для сохранения и введите имя файла в поле Файл, после чего нажмите кнопку Сохранить.

ДРУГИЕ СРЕДСТВА АВТОМАТИЗАЦИИ ПЕРЕВОДА

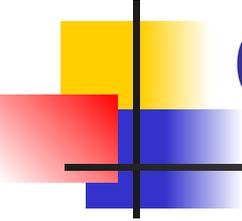


- Pragmatica;
- Word +.



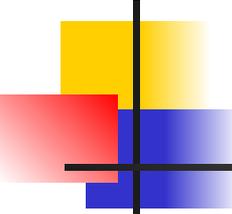
Пакет Pragmatica

- имеет собственную программную оболочку,
- машинный перевод осуществляется с использованием словаря общего назначения и ряда специализированных словарей,
- имеет возможность стыковки с текстовым редактором MS Word.



Система Word +

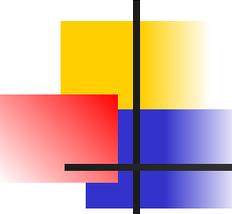
- работает со словарями меньшего объема,
- не имеет собственной оболочки, а встраивается прямо в текстовый редактор MS Word.

The logo graphic consists of a vertical black line intersected by a horizontal black line. To the left of the vertical line, there are three overlapping squares: a yellow one at the top, a red one in the middle, and a blue one at the bottom. The text 'ABBY Lingvo' is positioned to the right of the vertical line, with 'ABBY' in a bold blue font and 'Lingvo' in a blue font with a white outline.

ABBY Lingvo

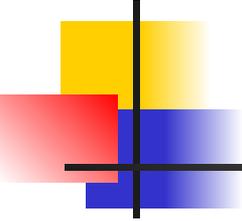
- Нажатие «горячей клавиши» в любом Windows-приложении за считанные секунды выведет на экран информацию, отобранную для перевода слова (фразы) из всех словарей, подключенных к системе.

ОСНОВНЫЕ ВОЗМОЖНОСТИ

The logo graphic consists of a vertical black line intersected by a horizontal black line. To the left of the vertical line, there are three overlapping squares: a yellow one at the top, a red one in the middle, and a blue one at the bottom. The word "Lingvo" is written in a blue, sans-serif font to the right of the vertical line.

Lingvo

- грамматические комментарии на любое слово,
- озвучивание наиболее употребляемых слов,
- проверка правильности написания,
- возможность создания собственных словарей.

- 
-
- Применение любого средства автоматизации машинного перевода значительно снижает время перевода и улучшает его качество.