Лабораторная работа 1 (уровень 1)

*Задания уровня 1.*

*1. 1*

Текст:

Проекты Cíbola/Oleada реализуют обширные компьютерные системы лингвистического анализа текстов, представленных в Unicode. Компоненты системы включают средства работы с мультиязыковыми текстами (MUTT), построения конкорданса (XConcord) для текстов на более чем 16 языках, статистического анализа, автоматического перевода, различные словари и тезаурусы. Некоторые версии этих компонентов доступны для бесплатной загрузки после процедуры формальной регистрации. Все компоненты реализованы в среде X11 Window System для SunOs и Solaris (источник: Проекты Cíbola/Oleada <http://rvb.ru/soft/catalogue/c01.html>).

|  |  |
| --- | --- |
| Слов | 67 |
| Символов (без пробелов) | 539 |
| Символов (с пробелами) | 605 |
| Символов в латинской графике | 95 |
| Чисел | 3 |
| Средняя длина слов | 8,04 |

Чтобы посчитать среднюю длину слов нужно количество символов (без пробелов разделить на количество слов.

*1.2*

Средняя длина слов русского языка: 7,9 (Количество русских символов: 444, количество русских слов: 56)

Средняя длина слов в латинской графике: 8,6 (Количество латинских символов: 95, количество латинских слов: 11)

Средняя длина текста: 8,04

Вывод: средняя длина слов текста соответствует русскому языку

*1.3*

а) Ø – норвежский, (ö ß) – немецкий, (ρ ω) – греческий, ё – русский, š – чешский.

б) th – английский, sch – немецкий, šč – чешский.

в) et – французский, the – английский, der – немецкий, och – шведский, için – турецкий.

*1.4*

Язык 1: русский язык

Язык 2: английский язык

|  |  |  |
| --- | --- | --- |
| Критерий | Язык 1: | Язык 2: |
| Типичные артикли | - | A, an, the |
| Указательные местоимения | Этот, тот, та, те, то, эта, эти, это, такой, таков | This, that, these, those |
| Местоимения 3-го лица | Он, она, оно | He, she, it |
| Отдельные формы вспомогательных глаголов | Быть | Do (does и did), be (am, is, are, was, were, being, been), have (has, having, had)  Модальные вспомогательные глаголы: will, shall, would, should, may, might, must, can, could |
| Основные предлоги и союзы | Предллоги:  Союзы: | On, in, at, near, over, under  And, or, as, that, but, if, before |
| Другие частотные слова | Год, человек, время, дело, жизнь, день, рука | Time, day, year, people, work |

*1.5*

Отрывок текста на русском языке (И.Бунин «Журавли»):

 Ясный и холодный день поздней осени, еду ровной рысцой по большой дороге. Блеск низкого солнца и пустых полей, осеннее безмолвное ожидание чего-то. Но вот вдали, за мной, слышен треск колёс. Прислушиваюсь – треск мелкий, быстрый, треск беговых дрожек.

(213 символов без пробелов)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица встречаемости букв в распространенных европейских языках | | | | | | | | | | | |
| Английский Немецкий Итальянский Французский Испанский | | | | | | | | | | Русский | |
| Буква | Встречаемость |
| Е | 12,86 | Е | 17,76 | Е | 19,18 | Е | 14,15 | I | 12,04 | о | 13,15 |
| Т | 9,72 | S | 8,23 | N | 10,20 | А | 12,90 | Е | 11,63 | е | 9,86 |
| А | 7,96 | А | 7,68 | I | 8,21 | О | 8,84 | А | 11,12 | с, н | 7,04 |
| I | 7,77 | N | 7,61 | S | 7,07 | S | 7,64 | О | 8,92 | л | 5,16 |
| N | 7,51 | Т | 7,30 | R | 7,01 | I | 7,01 | N | 7,68 | и, й | 4,69 |
| R | 7,03 | I | 7,23 | Т | 5,68 | R | 6,95 | Т | 7,07 | ы, д, к | 3,76 |

В таблице представлены данные о буквах русского алфавита, которые обладают самой большой встречаемостью на основе моих вычислений.

т(7), в(6), п,б,а,(5), з,г(4), х,ь,у,ш,м(3), ц,ж(2), я,ч,ё,ю(1) - количество каждой из оставшихся букв в представленном мною тексте, можно посчитать встречаемость каждой из них (количество буквы / количество символов \*100)

*1.6*

Русский                                                                         Эсперанто

зеленое дерево                                                                  verda arbo

старый человек                                                                 maljuna viro

хороший друг                                                                    bela amiko

 Морфологические диагностические показатели:

Существительные (ед.ч) оканчиваются на – о.

Прилагательные (ед.ч) оканчиваются на – а.

Существительные (мн.ч) оканчиваются на – j (деревья - arboj).

Прилагательные (мн.ч) – aj (зеленые - verdaj).

Окончания притяжательных прилагательных:

-a (мой, моя, мое)

-j (твой, твоя, твое)

-li (его, ее, его)

Окончания наречий:

-e (хорошо - bone)

Окончания глаголов:

-as (настоящее время)

-is (прошедшее время)

-os (будущее время)

-us (условное наклонение)

-u (повелительное наклонение)

Существует два падежа: именительный и винительный

Глаголы настоящего, прошедшего и будущего времени по лицам и числам не изменяются.

*2.1*

С точки зрения компьютерной семантики лишнее слово «липа»

С точки зрения компьютерной грамматики лишнее слово «ромашка»

*2.2*

Определения элементов синтаксических моделей:

a) AnNn ↔ N(A)nNg  (быстрое движение ↔ быстрота движения)

An = прилагательное

Nn = существительное (именительный падеж)

Ng = существительное (родительный падеж)

N(A)n = существительное, образованное от прилагательного

b) VNa ↔ N(V)nNg (прибавляю число ↔ прибавление числа)

V = глагол

Na = существительное

Ng = существительное

N(V)n = существительное, образованное от глагола

c) N1nN2g ↔ A(N2)nN1n (права автора ↔ авторские права)

N1n = существительное

N2g = существительное

A(N2)n = прилагательное, образованное от существительного

d) VвNa ↔ N(V)nвNa  (возвожу в степень ↔ возведение в степень)

V = глагол

вNa = существительное с предлогом "в"

N(V)n = существительное, образованное от глагола

e) DV ↔ A(D)nN(V)n (сильно желать ↔ сильное желание)

D = деепричастие

V = глагол

A(D)n = прилагательное, образованное от деепричастия

N(V)n = существительное, образованное от глагола

*2.3*

Визит врача ↔ врачебный визит N1nN2g ↔ A(N2)nN1n

Ароматный сад ↔ аромат сада AnNn ↔ N(A)nNg

Выхожу из дома ↔ выход из дома VизNg ↔ N(V)nизNg

*2.4*

а) N1n из N2g ↔ A(N2)nN1n

Костюм из шелка шелковый костюм.

Дом из кирпича кирпичный дом.

Мебель из дерева деревянная мебель.

б) AnNnVDAaNa

Красивая девочка нарисовала впереди зелёный автомобиль.