1. **Лингвистическая защита информации** (ЛЗИ) — это защита информации на этапе опредмечивания, овеществления смыслов, когда адресант, апеллируя к адресату, воплощает эти смыслы в сообщении. **Для защиты персональных данных лингвисты могут использовать следующие методы:**

* Анализ идентификационной информации.
* Анализ контекста.
* Разработка политик безопасности.
* Усовершенствование онлайн-инструментов для защиты персональных данных.
* Разработка правильной лингвистической тактики в социальных сетях и онлайн-мессенджерах.

1. Прикладная лингвистика не является синонимом терминов «компьютерная лингвистика», «вычислительная лингвистика» или «инженерная лингвистика», хотя в зарубежных исследованиях под термином «computational linguistics» понимаются все эти сферы и в научной литературе не встречаются отдельные термины для каждого из этих направлений. Нужно понимать, что прикладная лингвистика - это не использование компьютерных программ, это решение проблем, связанных с языком, это «приложение» любых знаний и технологий для решения практических задач.
2. Основными направлениями, рассматриваемыми в курсе компьютерной лингвистики, являются: информационный поиск, машинный перевод, терминология, терминоведение, терминография, компьютерная лексикография, распознавание и синтез речи, проблемы обучения языку с помощью компьютера. К компьютерной лингвистике относят все, что связано с использованием компьютеров в языкознании. Машинный перевод (Machine Translation) [20] — самое раннее приложение КЛ, вместе с которым возникла и развивалась сама эта область. Первые программы перевода были построены в середине прошлого века и были основаны на простейшей стратегии пословного перевода. Однако довольно быстро было осознано, что машинный перевод требует гораздо более полной лингвистической модели. Такая модель была разработана в отечественной системе ЭТАП [24], а также в нескольких других системах, выполняющих перевод научных текстов. В настоящее время существует целый спектр компьютерных систем машинного перевода (разного качества), от больших интернациональных исследовательских проектов до коммерческих автоматических переводчиков. Существенный интерес представляют проекты многоязыкового перевода с использованием промежуточного языка, на котором кодируется смысл переводимых фраз.

Упражнения: Уровень 1

Раздел 1:

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

1. Средняя длина слова составляет около 6 букв, но зависит от корпуса (набора) текстов. Для корпуса русской разговорной речи средняя длина слова обычно составит от 3.9 до 4.9 букв. Для корпуса русской художественной литературы - от 4.9 до 5.9. Для газетно-новостного-корпуса - от 5.9 до 6.9. И наконец, для корпуса научной и деловой литературы - от 6.9. Средняя длина английского слова составляет пять букв. Следовательно, средняя длина слов текста смешанного типа, представленного в задании 1, соответствует русскому языку.
2. а) Ø – датский, (ö ß) – немецкий, ρ ω – греческий, ё – русский, š – эстонский; б) th – английский, sch – немецкий, šč – словенский; в) et – французский, the – английский, der – немецкий, och – шведский, için – турецкий.

|  |  |  |
| --- | --- | --- |
| Критерий | Язык 1: Английский | Язык 2: Испанский |
| Типичные артикли | the, a, an | el, la, los, las |
| Указательные местоимения | this, that, these, those | este, esta, estos, estas, ese, esa, esos, esas |
| Местоимения 3-го лица | he, she, it | él, ella, ello |
| Отдельные формы вспомогательных глаголов | do, does, did; will, would; can, could | haber (he, has, ha, hemos, han); ser (soy, eres, es, somos, son); estar (estoy, estás, está, estamos, están) |
| Основные предлоги и союзы | in, on, at; and, but, or | en, a; y, pero |
| Другие частотные слова | to, is, are; not, can; I, you; he, she; we, they | de, que; no; yo, tú; él, ella; nosotros/nosotras, ellos/ellas |

1. Для дополнения таблицы встречаемости букв в распространенных европейских языках добавлю данные по русскому языку на основе следующего текста: "Русский язык - один из самых красивых и богатых языков мира. Он обладает богатой историей, разнообразным словарным запасом и глубокими корнями. В русском языке используется кириллический алфавит, который состоит из 33 букв. Среди самых часто встречающихся букв в русском языке можно выделить буквы: о, а, е, и, н, т, с, р, в, л." Теперь дополненная таблица встречаемости букв в распространенных европейских языках выглядит следующим образом:

|  |  |  |  |
| --- | --- | --- | --- |
| Буква | Английский язык | Испанский язык | Русский язык |
| А | 8,2% | 12,5% | 7,6% |
| Б | 1,5% | 0,9% | 1,6% |
| В | 3,3% | 1,4% | 4,3% |
| Г | 1,9% | 1,1% | 1,7% |
| Д | 4,2% | 3,1% | 2,9% |
| Е | 12,7% | 14,7% | 8,4% |
| Ж | 0,6% | 0,4% | 0,8% |
| З | 1,7% | 0,7% | 1,6% |
| И | 7,5% | 7,5% | 6,2% |
| Й | 1,5% | 0,5% | 1,2% |
| К | 6,6% | 3,3% | 3,5% |
| Л | 4,3% | 4,6% | 3,2% |
| М | 2,5% | 3,1% | 2,9% |
| Н | 7,2% | 7,5% | 6,7% |
| О | 7,6% | 5,4% | 10,9% |
| П | 2,0% | 3,1% | 2,5% |
| Р | 6,7% | 6,6% | 4,7% |
| С | 6,7% | 8,0% | 5,9% |
| Т | 9,1% | 7,3% | 6,8% |
| У | 2,8% | 4,4% | 2,6% |
| Ф | 0,2% | 0,9% | 0,3% |
| Х | 1,5% | 0,8% | 0,9% |
| Ц | 0,6% | 1,4% | 0,5% |
| Ч | 1,2% | 0,8% | 1,6% |
| Ш | 0,9% | - | 0,8% |
| Щ | 0,3$ | - | - |
| Ъ | - | - | - |
| Ы | - | - | - |
| Ь | - | - | - |
| Э | - | - | - |
| Ю | - | - | - |
| Я | - | - | - |

Теперь таблица содержит данные по встречаемости букв в английском, испанском и русском языках для сравнения и анализа.

1. Морфологические диагностические показатели эсперанто: 1. **Суффиксы**: В эсперанто существительные и прилагательные имеют одинаковые окончания. Например, суффикс "-a" обозначает прилагательное (verd**a** - зелен**ый**), а суффикс "-o" обозначает существительное (arb**o** - дерев**о**). 2. **Грамматическое родоведение**: В эсперанто нет грамматического рода у существительных и прилагательных. Они не изменяются по роду, что упрощает грамматическую структуру языка. 3. **Единая форма для разных частей речи**: В эсперанто одинаковые окончания используются для образования различных частей речи, что делает язык более предсказуемым и легким для изучения. 4. **Словообразование**: В эсперанто слова образуются посредством аффиксации (добавление приставок и суффиксов) к корням, что делает процесс формирования новых слов структурированным и прозрачным.

Раздел 2:

1. С точки зрения компьютерной семантики и компьютерной грамматики можно рассмотреть данную задачу следующим образом: 1. **Ландыш** - это цветок. 2. **Левкой** - это также цветок. 3. **Лаватера** - это цветок. 4. **Лютик** - это цветок. 5. **Люпин** - это цветок. 6. **Ромашка** - это цветок. 7. **Липа** - это дерево. Слово **Липа** является лишним в данном списке, так как все остальные слова представляют собой названия цветов, а "Липа" - это дерево.
2. а) AnNn ↔ N(A)nNg (быстрое движение ↔ быстрота движения) - An - прилагательное, обозначающее характеристику или свойство - Nn - существительное, обозначающее объект или явление - Ng - существительное, обозначающее абстрактное понятие - N(A)n - существительное, обозначающее объект или явление, к которому относится прилагательное - Пример цитаты: "быстрое движение ↔ быстрота движения" b) VNa ↔ N(V)nNg (прибавляю число ↔ прибавление числа) - V - глагол, обозначающий действие - Na - существительное, обозначающее объект или явление - N(V)n - существительное, обозначающее объект или явление, к которому относится глагол - Ng - существительное, обозначающее абстрактное понятие - Пример цитаты: "прибавляю число ↔ прибавление числа" c) N1nN2g ↔ A(N2)nN1n (права автора ↔ авторские права) - N1 - существительное, обозначающее объект или явление - N2 - существительное, обозначающее объект или явление - Ng - существительное, обозначающее абстрактное понятие - A(N2)n - прилагательное, обозначающее характеристику или свойство, к которому относится существительное - Пример цитаты: "права автора ↔ авторские права" d) VвNa ↔ N(V)nвNa (возвожу в степень ↔ возведение в степень) - V - глагол, обозначающий действие - Na - существительное, обозначающее объект или явление - N(V)n - существительное, обозначающее объект или явление, к которому относится глагол - в - предлог - Пример цитаты: "возвожу в степень ↔ возведение в степень" e) DV ↔ A(D)nN(V)n (сильно желать ↔ сильное желание) - D - наречие, обозначающее степень действия - A - прилагательное, обозначающее характеристику или свойство - (D)n - наречие, обозначающее степень действия, к которому относится прилагательное - N(V)n - существительное, обозначающее объект или явление, к которому относится глагол - Пример цитаты: "сильно желать ↔ сильное желание"
3. а) AnNn ↔ N(A)nNg - визит врача ↔ врачебный визит - Пример цитаты: "визит врача ↔ врачебный визит" б) N1nN2g ↔ A(N2)nN1n - ароматный сад ↔ аромат сада - Пример цитаты: "ароматный сад ↔ аромат сада" в) VNa ↔ N(V)nNg - выхожу из дома ↔ выход из дома - Пример цитаты: "выхожу из дома ↔ выход из дома"
4. а) книга из библиотеки ↔ библиотечная книга. Пример цитаты: "Я взял книгу из библиотеки и узнал, что это библиотечная книга." б) зеленая трава шумно шелестит под легким ветром.