**Выявление ориентиров для перевода французского неопределенно-личного местоимения «on» с помощью модели искусственного интеллекта**

***Шлычкова Елизавета Олеговна***

*Студент*

*Тюменский государственный университет,*

*институт социально-гуманитарных наук, Тюмень, Россия*

*E–mail: E3286@yandex.ru*

Перевод французского местоимения «on» имеет свои особенности ввиду многообразия выражаемых им значений, а потому представляет сложность как для переводчика-человека, так и для систем машинного перевода. В данном исследовании предлагается использование модели искусственного интеллекта – многослойного перцептрона, для выявления наиболее значимых аспектов, позволяющих подобрать точный эквивалент местоимения «on» в языке перевода. Объектом нашего исследования является местоимение «on» во французском языке и его варианты перевода на русский язык. Предмет исследования – контекстуальные ориентиры, позволяющие подобрать верный эквивалент местоимения в языке перевода. Цель работы – отбор наиболее значимых аспектов контекста с помощью многослойного перцептрона и генетического метода отбора признаков.

В качестве материала исследования были отобраны 3620 пар фраз из корпуса параллельных текстов [4], в которых содержится местоимение «on». Для каждого слова в отобранном материале были определены часть речи и некоторые грамматические признаки на основе алгоритма, представленного в одной из наших работ «Система правил для машинного перевода русских местоимений второго лица на французский язык» [2]. Следующим этапом было выявление способов перевода местоимения в оригинальном корпусе. По словам Е. А. Рощупкиной, «конструкция с **on** соответствует в русском языке неопределенно-личной форме», а также «**on** может относиться в речи к определенным лицам и иметь в зависимости от ситуаций значение любого лица» [1], что мы и взяли за основу при определении вариантов перевода местоимения. Статистические данные распределения вариантов перевода представлены в таблице 1.

*Таблица 1. Количество переводов для каждого из вариантов*

|  |  |  |  |
| --- | --- | --- | --- |
| **Глагол 3 л. ед. ч.** | 837 | **Мы** | 72 |
| **Он** | 416 | **Глагол 3 л. мн. ч.** | 34 |
| **Я** | 225 | **Ты** | 30 |
| **Она** | 171 | **Оно** | 16 |
| **Они** | 149 | **Неизвестно** | 1586 |
| **Вы** | 84 | **Всего:** | **3620** |

«Неизвестные» переводы могут являться следствием ошибочного автоматизированного отбора пар фраз без соответствующего местоимения, либо наличия перевода с некоторым отклонением от оригинальной фразы (например, исключения фразы с данным местоимением при переводе). Данные пары фраз с неизвестным переводом были исключены из итоговой выборки.

Для сравнения работы многослойного перцептрона и существующих систем машинного перевода, мы получили результаты перевода половины пар фраз корпуса популярными по результатам поисковых систем на 21.02.2021 онлайн-переводчиками: «Яндекс.Переводчик», «Google Переводчик» и «Reverso». Средняя точность перевода составила 0,29. Стоит заметить, что полученные данные не говорят об истинной точности перевода, так как у местоимения «on» нет однозначного эквивалента – данные результаты необходимы для сравнения точности многослойного перцептрона при отборе ориентиров для перевода местоимения «on».

В качестве признаков для обучения многослойного перцептрона были выбраны характеристики окружения (в радиусе трех элементов) местоимения «on», а именно: часть речи, лицо, число, а также тип предложения. Целевой признак – вариант перевода местоимения «on». В результате точность многослойного перцептрона достигла 0,33. Чтобы выявить наиболее значимые ориентиры при переводе данного местоимения, мы обратились к генетическому методу отбора признаков, более подробный алгоритм которого представлен в работе «Предсказание операторов внутренней речи на основе данных ЭЭГ с помощью метода ближайших соседей» [3]. Данный алгоритм позволил повысить точность предсказания до 0,4. Значимыми для перевода местоимения «on», с точки зрения модели, оказались признаки-ориентиры, представленные на рисунке 2.

*Рисунок 2. Отобранные признаки-ориентиры для перевода местоимения «on»*

Число

-

Лицо

**On**

Лицо

-

-

**1 2 3 4 5 6**

Мы можем прийти к выводу, что большую роль при выборе способа перевода местоимения «on» играет лицо элемента, стоящего до или после местоимения. Так как с местоимением «on» глагол употребляется в форме третьего лица единственного числа, вероятно, что модель ориентировалась на лицо местоимения, играющего роль дополнения, или на косвенное наличие инверсии. Примечательным является включение такого признака, как число первого элемента. Чаще всего на месте второго элемента находился знак пунктуации, а потому первый элемент может являться референтом местоимения «on», что обосновывает необходимость акцентирования на нем внимания при выборе способа перевода. Таким образом, выявленные признаки могут во многом способствовать усовершенствованию систем машинного перевода при интерпретации местоимения «on», а выработанный алгоритм может в перспективе использоваться для анализа признаков при переводе иных сложных лексических и грамматических явлений. Более того, представленные данные могут применяться и при «ручном» переводе многофункционального местоимения «on».

**Литература**

1. Рощупкина Е.А. Краткий справочник по грамматике французского языка. / Е.А. Рощупкина - М.: Высшая школа, 1990. 239 с.
2. Шлычкова Е.О. Система правил для машинного перевода русских местоимений второго лица на французский язык / Е.О. Шлычкова // Материалы Международного молодежного научного форума «ЛОМОНОСОВ-2023». - 2023. URL: https://lomonosov-msu.ru/archive/Lomonosov\_2023/data/section\_38\_28603.htm (дата обращения: 11.08.2023)
3. Шлычкова Е.О. Предсказание операторов внутренней речи на основе данных ЭЭГ с помощью метода ближайших соседей / Е.О. Шлычкова, А.Н. Шевляков // Математическое и информационное моделирование : материалы Всероссийской конференции молодых ученых. - 2023. - С. 116-124.
4. Параллельные тексты на французском и русском языках: http://lingvo.asu.ru/france/texts\_fr.html