

В. И. Керножицкая (Минск, МГЛУ)

ПРИЧИНЫ ВОЗНИКНОВЕНИЯ ОТКЛОНЕНИЙ ОТ ЯЗЫКОВЫХ НОРМ В СООБЩЕНИЯХ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ

В статье рассматриваются два вида факторов, влияющих на появление отклонений от языковых норм в виртуальном общении пользователей социальных сетей. Подчеркивается, что первый фактор – человеческий – связан со свойственным людям стремлением к экономии языковых средств в процессе любого общения, а также обусловлен спонтанностью, неподготовленностью высказываний виртуальных коммуникантов, выражающейся в вербальной компрессии на разных языковых уровнях. Отмечается возникновение своеобразной компенсации прямого невербального общения путем использования компактных графических средств. Второй фактор появления отклонений от норм литературного языка обусловлен спецификой интернет-среды. Готовые единицы языка переносятся в интернет-коммуникацию, где частично трансформируются под воздействием законов и правил сетевого общения. С целью дальнейшей автоматической обработки высказываний пользователей социальных сетей встречающиеся в них ненормализованные языковые единицы необходимо предварительно привести к литературной норме, используя различные методы.

Новые возможности взаимодействия пользователей в Интернете формируют новые стандарты виртуального общения. Среди особенностей интернет-коммуникации, прежде всего, необходимо выделить тенденцию к использованию более свободной формы высказывания и более свободному употреблению языковых средств, вплоть до отклонения от норм литературного языка. Специалисты отмечают, что источником появления отклонений от норм языка могут быть как люди, так и среда, в которой они общаются. В аспекте указанной проблемы речь в статье пойдет о такой среде интернет-общения, как социальные сети, а также о пользователях социальных сетей.

Необходимо отметить, что, в первую очередь, на изменения в языке, выраженные отклонением от языковых норм, влияют сами люди. Одна из наиболее явных причин этого явления – стремление пользователей социальных сетей к упрощению языковых средств, т.е. к их компрессии. Механизм данного процесса предопределен законом экономии языковых средств, когда одна языковая единица осуществляет функцию нескольких языковых единиц, т.е. «в одном означающем слито несколько означаемых» [1, с. 605]. Происходит это по причине того, что человеческому мышлению и, соответственно, языку, свойственно стремление к экономии усилий и ресурсов, суммированию и обобщению отражаемых событий. При описании различных жизненных ситуаций «элементарные мысли и выражающие их элементарные предложения могут тем или иным способом сжиматься или перекрываться» [2, с. 263]. Языковая компрессия определяется самой природой человеческого мышления и временем. Невозможно каждый раз передавать с помощью языка то, что происходило в реальности, до мельчайших подробностей, да и вряд ли это нужно [2, с. 264]. Особенно ярко языковая компрессия выражается в условиях интернет-общения, которое подразумевает ограниченность коммуникативных ресурсов во времени реакции

на высказывание собеседника (если речь идет о синхронном общении), в объеме сообщения, в графических средствах репрезентации, в использовании невербальных средств коммуникации [3, с. 83]. Одной из причин является и то, что большая часть текстов, порождаемых пользователями социальных сетей, представляет собой спонтанные, неподготовленные высказывания [4].

Таким образом, отклонения от норм языка могут осуществляться как стихийно, так и осознанно. К нарушениям, возникающим стихийно, относятся разного рода опечатки, к которым в сетевой коммуникации относятся весьма снисходительно, а также пренебрежение правилами орфографии, обусловленное ситуацией общения в условиях ограниченности коммуникативных ресурсов. К осознанным нарушениям относится употребление эрративов – специальных искажений стандартного написания слов носителями языка, владеющими литературной нормой. Распространение эрративов демонстрирует высокую степень восприимчивости интернет-коммуникации к новым, нестандартным, в том числе девиантным языковым явлениям, если они достаточно выразительны. Кроме того, в общесетевом языке широко распространены эрративы, образованные по принципу «как слышится – так и пишется», которые употребляются для создания эффекта спонтанности, неподготовленности и непринужденности разговорной речи. В электронной среде происходят отклонения как в сторону фонетического письма, когда язык как бы «привязывается» к звучанию речи, так и в сторону идеографического письма (к которому можно отнести эмодзи и аббревиатуры, по сути являющиеся визуальными шифрами, например, *LOL – laughing out loud*, *abt – about* и т.п.) [3, с. 85].

Еще одним фактором, влияющим на появление отклонений от норм в языке пользователей социальных сетей, является так называемая «компенсация» невербального общения или недостаток экспрессии в выражении своих чувств и эмоций по тому или иному поводу, как при живом общении. Это происходит потому, что, в отличие от традиционной устной речи, в сети не устанавливается мгновенная обратная связь между участниками в форме сигналов непосредственной реактивной (речевой или неречевой) реакции собеседника. Даже в таких наиболее интерактивных форматах интернет-коммуникации, как чаты и общение, через службу мгновенных сообщений, ритм обмена сообщениями отличается от ритма разговора [3, с. 84]. Так называемый «эмоциональный дефицит» компенсируется путем введения в виртуальную среду общения эмодзи (смайлов) и эмоджи – графических знаков выражения эмоций [4]. Они используются в целях репрезентации конкретных элементов языка мимики и жестов, а также определенных действий участников коммуникации (например, дарение цветка). Кроме того, различные эмоции (например, радость, удивление, шок) выражаются путем повторения гласных или согласных несколько раз в зависимости от характера и степени эмоции, например, *OOOO!* [3, с. 86]. Аналогичную роль могут выполнять пробелы, знак «*», нестандартные комбинации знаков препинания или акцентирование части фразы написанием заглавными буквами [4].

Другим источником появления отклонений от языковых норм в сообщениях пользователей социальных сетей является интернет-среда. Ее влияние на язык ощущается на всех уровнях языковой системы. Прежде всего, оно фиксируется на уровне лексики, а также затрагивает словообразовательные модели, нормы пунктуации и орфографии, особенности синтаксической структуры,

способы организации и структурирования текста и т.п. [3, с. 87]. Это происходит потому, что язык, даже тот, который используется в рамках определенного сообщества, с течением времени подвергается изменениям, адаптируясь под языковые нужды тех, кто на нем говорит [5, р. 9]. Любой язык адаптируется к потребностям общества, социальных групп и людей, которые его используют, но особенно ярко это проявляется в языке интернет-общения. Так большинство пользователей социальных сетей стремятся описать свои впечатления и эмоции, выразить свою точку зрения, используя характерные особенности интернет-коммуникации, но мало обращая внимание на нормы используемого языка.

В своем стремлении пользователи часто выходят за рамки существующих понятий и тем самым образуют новые пласты лексики. Активность пополнения языкового состава в интернет-среде несопоставима с реальной коммуникацией, так как процесс появления новых слов в языке виртуального общения достаточно прост и прямолинеен. Языковая единица не проходит долгий путь образования перед своим проникновением в словари и сферу употребления. Уже готовое языковое явление переносится напрямую в интернет-коммуникацию, где подвергается дальнейшей адаптации. Конечно, большинство новых лексических единиц не используется в течение длительного периода времени, но они оказывают определенное влияние на формирование языковых механизмов в целом. Ученые полагают, что возможно, в будущем данный процесс в какой-то мере утратит популярность и пойдет на спад, но в настоящее время человечество находится под активным воздействием информационных технологий, что влечет за собой ощутимые изменения в языковой среде [3, с. 88].

Таким образом, языковая картина общения в социальных сетях постоянно претерпевает изменения, которые необходимо тщательно исследовать, чтобы эффективно подходить к процессу нормализации языковых единиц в целях их дальнейшей автоматической обработки. Рассмотрим подробнее основные подходы к автоматической нормализации лексических единиц на примере обработки сообщений англоязычных пользователей социальной сети *Twitter*. Так, разработан метод нормализации твитов, совмещающий в себе как методы машинного обучения, так и правила [5, р. 11–12]. На начальном этапе происходит отбор всех некорректных слов с помощью модели машинного обучения, натренированной на их поиск по определенным атрибутам. Затем к каждому слову применяется одно из следующих правил нормализации: 1) расшифровка аббревиатур (при этом используется список самых частотных аббревиатур в сети *Twitter* и их расшифровок); 2) обработка причастий настоящего времени (пользователи часто пропускают в них букву *i* или *g*, например, употребляют *goin* вместо *going*. Данное правило направлено на устранение таких опечаток: сначала определяется часть речи, и если это причастие, то восстанавливается его нормальная форма); 3) обработка пропущенного апострофа (правило направлено на исправление ошибок в таких случаях, как *'m*, *'ll*, *'ve*, *'re*, *n't*, *'s*); 4) обработка повторяющихся символов (в данном случае происходит удаление избыточного количества повторяющихся символов). Группа исследователей представила еще один подход под названием *TENOR (TEText NORmalisation Approach)* [5, р. 13–14]. В данном случае процесс нормализации состоит из двух этапов – отбора не соответствующих нормам слов с помощью словаря и их дальнейшей замены на стандартную форму. Процесс замены включает в себя следующие действия: все

нестандартные символы и пунктуация, за исключением смайлов, удаляются с помощью регулярных выражений, затем расшифровываются аббревиатуры, смайлы, сокращения и устраняются опечатки. Разработана еще одна система лексической нормализации, состоящая из двух блоков [5, p. 15]. В первом блоке происходит выделение потенциальных кандидатов для нормализации с использованием метода *CRF* (*Conditional Random Fields*, разновидность метода марковских случайных полей), натренированного на корпусе из 2 950 размеченных твитов, а во втором блоке – их последующая нормализация. Стратегия лексической нормализации твитов, предложенная еще одной группой разработчиков, включает в себя следующие шаги: 1) генерацию набора слов, являющихся потенциальными кандидатами на нормализацию; 2) определение всех возможных вариантов для нормализации каждого слова, отбор наиболее вероятного варианта путем сравнения с помощью алгоритмов, которые позволяют, например, в зависимости от контекста привести слово *mve* к слову *me* или к слову *move*; 3) приведение слов к их стандартной форме [5, p. 16]. Таким образом, в общем плане процедура автоматической нормализации лексических единиц осуществляется в два этапа. На первом этапе происходит отбор не соответствующих лексическим нормам слов твита, а на втором – их последующая нормализация либо на основе стандартных алгоритмов, либо методами машинного обучения.

ЛИТЕРАТУРА

1. Кибрик, А. Е. Язык / А. Е. Кибрик // Лингвистический энциклопедический словарь. – М., 1990. – С. 604–606.
2. Умерова, М. В. Языковая компрессия: виды и уровни реализации / М. В. Умерова // Актуальные вопросы современной науки. – 2011. – №17-(1). – С. 260–269.
3. Максимова, О. Б. Язык в интернет-коммуникации: общие закономерности и национально-культурные особенности (на материале русского и английского языков) / О. Б. Максимова // Вестник РУДН. Сер. Теория языка. Семиотика. Семантика. – 2010. – № 3. – С. 83–89.
4. Тюленева, В. Н. Язык Интернета: характеристика, особенности и влияние на речь / В. Н. Тюленева, И. А. Шушарина // Вестн. Курган. гос. ун-та. – 2018. – № 1 (48) [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/yazyk-interneta-harakteristika-osobennosti-i-vliyanie-na-rech>. – Дата доступа: 28.04.2019.
5. Reitter, D. Did Social Networks Shape Language Evolution? / D. Reitter, Ch. Lebiere // Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics, Uppsala, 15 July 2010. – Sweden, 2010. – P. 9–17.

The article deals with the problem of deviations from language norms in virtual communication of social networks users. It's emphasized that the first reason of this phenomenon is the human one which leads to verbal compression at different language levels. The second factor is due to the specifics of Internet communication.