

Морфосинтаксис полипредикации в устной речи в сравнении с письменной по данным корпуса японских личных рассказов¹

Подлесская В.И. (РГГУ\РАНХиГС); Озолина В.А. (РГГУ)

В работе ставится следующий исследовательский вопрос: различается ли частотность и номенклатура морфосинтаксически маркированных межпредикатных связей в устном и письменном дискурсе?

В рамках современных холистических подходов к оценке языковой сложности (Miestamo 2017) синтаксическую сложность нередко измеряют с помощью мер полипредикативности, выработанных в прикладной лингвистике для аттестации уровня освоения (родного или изучаемого) языка. Эти меры, так или иначе, работают с понятием полипредикативного комплекса (ППК) – «сложного предложения» или его аналогов (ср. «T[erminable]-unit», Hunt 1965, Young 1995, Ortega 2003, 2012; «clause-package», Berman, & Nir-Sagiv 2009, Berman 2016; «A[nalysis of] S[peech]-Unit», Foster et al. 2000) и оценивают (1) объем ППК, т.е. количество клауз в его составе; и (2) плотность полипредикации, т.е. долю клауз, входящих в ППК, относительно общего объема текста.

В исследованиях по устной речи неоднократно высказывалось соображение, что в устном тексте основные меры полипредикативности ниже, чем в письменном (Земская и др. 1981, Сиротинина 1974, Лаптева 1976). Однако в отсутствие фронтальной эмпирической проверки к такого рода утверждению надо подходить осторожно по следующим причинам.

Первое. Меры полипредикативности существенно варьируются в зависимости от того, как в конкретном исследовании оценивается вклад просодии в организацию полипредикации, Кибрик 2008. Например, по-разному интерпретируется пара клауз, таких, что семантическая связь между ними маркируется только просодически, без подключения морфосинтаксических средств («бессоюзная связь»). Если мы принимаем в расчет просодию, то такая пара клауз будет идентифицироваться как ППК, если нет – то это две изолированных клаузы.

Второе. Тезис о низких мерах полипредикативности имплицитно, а иногда и эксплицитно относится не к устной речи вообще, а к разговорной речи и даже уже – к неподготовленному непубличному диалогическому дискурсу (Земская и др. 1981, Biber 1988, Miller 1995, Miller & Fernandez-Vest 2006, Miller & Weinert. 2009 [1998], Haselow 2017 и многие другие). Поэтому, сравнивая существующие меры полипредикативности, необходимо делать поправки на то, какие типы дискурса используются авторами в качестве эмпирической базы.

Третье. Далеко не очевидно, что сравнительные оценки мер полипредикативности отражают универсальные тенденции, а не лингвоспецифические феномены. Так, в ряде исследований отмечалась такая особенность устной японской речи, как (сверх) длинные полипредикативные комплексы с грамматически маркированными межфразовыми связями (Clancy 1982, Iwasaki & Ono 2001, Suzuki 2013, Laurý & Ono 2014) – что, по-видимому, не характерно для языков «среднеевропейского стандарта». В этом отношении японский язык проявляет свойство, типичное для языков левого ветвления, активно эксплуатирующих феномен clause-chaining, Longacre 1990, Payne 1991. Можно предположить, что механизм clause-chaining в таких языках выполняет ту же функцию, что и последовательность просодически незавершенных клауз, связь между которыми может не маркироваться морфосинтаксически.

Учитывая все сказанное, мы решили провести эмпирическую проверку мер полипредикативности таким образом, чтобы противопоставление по каналу передачи

¹ Работа поддержана РФФИ, грант 16-06-00226

(устный vs. письменный) оказалось единственной переменной, по которой противопоставляются две группы текстов, а остальные признаки – жанр, язык, набор испытуемых, способ элицитации – были бы закреплены. В качестве материала мы использовали специально собранный нами корпус устных неподготовленных личных рассказов на японском языке. В первую сессию эксперимента испытуемым ($n=14$) было предложено рассказать веселую историю из своего личного опыта. Во вторую сессию (через сутки) тем же испытуемым было предложено самостоятельно записать ту же историю. Таким образом, мы получили 14 пар рассказов, представленных в двух версиях – устной (в виде аудиофайла и транскрипта) и авторской письменной. До начала обеих сессий предполагаемое задание испытуемым известно не было, что повысило меру спонтанности элицированных нарративов. Все рассказы были сегментированы: были выделены все клаузы и все сегменты, не инкорпорированные в клаузальную структуру (дислоцированные топики, отделенные от своей базовой клаузы другим материалом; автономные дискурсивные маркеры; «обломки» клаузальной структуры, возникшие в результате речевого сбоя, и ряд других). Каждая клауза была проаннотирована либо как изолированная, либо как связанная, т.е. входящая в ППК с соответствующим статусом (главная клауза, адвербиальная, относительная или сентенциальный актанта; при многоуровневом вложении одна и та же клауза могла получить комплексный статус, будучи одновременно, например, главной по отношению к вложенной и актанта по отношению к более «высокой» главной). Вхождение в ППК обеспечивалось только морфосинтаксическим маркированием (с помощью особого порядка слов, инфинитных форм, а также сочетаний финитных и инфинитных форм с коннекторами – союзами, субстантиваторами, послелогоми); связи, маркированные только просодически, не учитывались. Каждый маркер был проаннотирован, и таким образом была определена полная номенклатура средств межфразовой связи, зафиксированных в корпусе. Далее, для всех текстов корпуса были вычислены основные метрики полипредикативности: число связанных клауз (входящих в ППК), число изолированных клауз, число неклаузальных сегментов, число ППК, объем ППК (число связанных клауз к числу ППК), плотность полипредикации (доля связанных клауз от общего числа клауз), частотность отдельных маркеров полипредикативности (в приведении к числу клауз в тексте) и ряд других метрик. По этим метрикам мы сравнили медианные значения в устных и письменных рассказах с применением непараметрического статистического критерия Уилкоксона. Наиболее существенные количественные результаты сводятся к следующему.

- (1) Плотность полипредикации не демонстрирует значимых отличий между устными и письменными текстами ($p>0.05$) – $Me_{уст} = 0.9715$; $Me_{письм} = 0.9665$
- (2) В устных текстах значимо выше ($P_{1-tail} = 0.0005$, $P_{2-tail} = 0.001$):
 - (а) объем ППК – $Me_{уст} = 9.875$; $Me_{письм} = 3.9545$;
 - (б) суммарная частотность основных сочинительных союзов *ga* и *ke(re)do(mo)* в приведении к числу клауз – $Me_{уст} = 0.111$; $Me_{письм} = 0.0405$;
 - (в) частотность основных нарративных конвербов (формы на *-i* и на *-te*) в приведении к числу клауз – $Me_{уст} = 0.2825$; $Me_{письм} = 0.1835$

Таким образом, утверждение о меньшей полипредикативности устных текстов не прошло проверку на универсальность. Несмотря на то, что в устных рассказах у говорящих имеется такой мощный дополнительный ресурс маркирования полипредикации, как просодия, они, как оказывается, используют морфосинтаксическое маркирование с не меньшей интенсивностью, чем в письменных рассказах. Разумеется, данные получены для конкретного языка с выраженным левым ветвлением и конкретного типа дискурса – устного нарратива. Дальнейшие исследования в этом направлении помогут проследить связь метрик полипредикативности с типологическим статусом языка, жанром дискурса и другими параметрами.