

Научная статья
УДК 81-139
DOI 10.18101/978-5-9793-1802-8-2022-179-185

ИССЛЕДОВАНИЕ ФУНКЦИЙ ФРАЗОВЫХ ЧАСТИЦ КИТАЙСКОГО ЯЗЫКА НА МАТЕРИАЛЕ ЛИНГВИСТИЧЕСКОГО КОРПУСА

© Павлова Алена Дмитриевна

кандидат филологических наук, преподаватель,
Военный университет имени князя Александра Невского МО РФ
Россия, г. Москва
ahisha@mail.ru

Аннотация. Роль фразовых частиц в построении предложений уникальна. Они могут выполнять риторические, экспрессивные, анафорические, грамматические функции. В связи с этим перевод фразовых частиц в парах языков разной типологии, где функции фразовых частиц оригинала могут быть переданы разными средствами языка перевода (интонационно, грамматически, лексически и т. д.), представляется достаточно сложным. Типология китайского языка как изолирующего, в котором роль служебных слов значительнее в выражении синтаксических отношений между словами, обуславливает обязательность и большую частотность использования фразовых частиц для построения предложений, выражения отношения между его частями.

В статье рассматриваются подходы к исследованию функций фразовых частиц китайского языка. По результатам проведенного исследования предложен алгоритм анализа функций фразовых частиц, описаны преимущества и особенности его проведения на материале лингвистического корпуса.

Ключевые слова: фразовая частица, лингвистический корпус, модальные частицы, методы исследования, китайский язык.

CHINESE DISCOURSE PARTICLES' LINGUISTIC CORPUS BASED ANALYSIS

© Alena D. Pavlova

PhD, Senior Lecturer,
Military University Ministry of Defense of the Russian Federation
Russia, Moscow
ahisha@mail.ru

Abstract. The role of sentence-final particles (discourse particles) in the construction of sentences is unique. They can perform rhetorical, expressive, anaphoric, grammatical functions. In this regard, the translation of phrasal particles in pairs of languages of different typologies, where the functions of the original sentence-final particles can be conveyed by different means of the target language (by intonation, grammatically, lexically, etc.), seems to be quite difficult. The typology of the Chinese language make the role of function words more significant in expressing the syntactic relations between words, determines the obligatoriness and greater frequency of using sentence-final particles to build sentences, expressing the relationship between its parts. This article is dedicated to the *research methods* of the Chinese sentence-final particles (discourse particles) functions.

Based on the linguistic corpus research of the discourse particles function we propose the analyzing algorithm, describe the advantages and difficulties of the corpus based methods.

Keywords: sentence-final particles, discourse particles, linguistic corpus, research methods, Chinese language.

Термин «фразовые частицы» в отечественном китаеведении был впервые предложен В.И.Гореловым в книге «Грамматика китайского языка» (1982) [1]. В работах российских исследователей также используются термины «модальная частица», «фразовая модальная частица». В работах западных исследователей наиболее современными являются термины *discourse/pragmatic particle* (*sentence-final particle* – все еще довольно частотный термин, однако уже признан устаревшим в отношении китайского языка, как не отражающий функциональные особенности частиц). В исследованиях современного китайского языка китайскими лингвистами фразовые частицы относят к самостоятельной категории служебных слов и называют 语气词 (дословно «модальные слова») (Ван Ли, Ма Чжэнь, Чжу Дэси, Хуан Божун, Чжан Чжигун, Ху Юйшу, Ци Хуъян, Хоу Сюэчао, Сунь Синь, Го Силян, Ли Сяоцзюнь и т. д.) [6].

Исследования фразовых частиц следуют за развитием лингвистической науки. В отечественной лингвистике фразовые частицы рассматриваются с точки зрения их позиции в предложении и структуры (В. И. Горелов) [1], пресуппозиции и импликации (Тань Аошуан) [8], выражения оттенков экспрессивно-модального отношения говорящего и позиции в предложении (Е. И. Шутова) [9], указания на изменение характера согласования деятельности между коммуникантами при тождественном наполнении предложения (В.А. Курдюмов) [3]. В работах западных синологов основными подходами к исследованию фразовых частиц нами выделены функциональный (Чжао Юаньжэнь, Ч. Ли и С. Томпсон) [10;12] и дискурсивный подходы (Ruey -Jinan Regina Wu, Li Boya) [11;13].

Однако в работах теоретического и справочного характера по исследованию функций фразовых частиц как китайских, так и отечественных, западных исследователей анализ проводится на языковом материале преимущественно в виде отдельных фраз или предложений. При перечислении функций предлагаются отдельные примеры для подтверждения определенной функции. Как правило, примеры использованы из литературных произведений или придуманы автором [14;1;3].

Рассмотрим на примере 1 фразы «Ты ученик» с использованием разных фразовых частиц в конце.

Пример 1.

你是学生吗?

你是学生吧。/?!

你是学生呢? /。

你是学生啊! /?/。

你是学生了。

Перед нами несколько предложений, которые отличаются только фразовой частицей в конце. Можем ли мы дать точные переводы этим выражениям? Для

предложений с частицами 吗, 吧 смысл предложения достаточно очевиден: вопросительное предложение с частицей 吗, подразумевающее, что ответ будет да или нет; с частицей 吧 в зависимости от типа предложения (повествовательное, вопросительное, восклицательное) значение будет немного отличаться, но в любом случае частица привнесет значение почти стопроцентной уверенности говорящего в том, что «ты ученик». Однако если бы не было знаков препинания у частиц 吧, 呢, 啊, мы не смогли бы восстановить интонацию и дать точный перевод фразы. Также неоднозначна частица 了, так как может быть и утвердительной частицей, и показателем изменения ситуации, а также совмещать эти два значения. И без контекста «дешифровать» такие предложения крайне затруднительно.

Контекст фразы, отношение говорящего и скрытая семантика, выраженная в частице, часто не поддаются «дешифровке» только на основании одного предложения. Наши лингвистические эксперименты с носителями по восстановлению фразовых частиц в тексте подтверждают, что фразовая частица является маркером модальности, который дополнительно фиксирует отношение говорящего к контексту, интонацию, и является часто вторичной, усиливающей уже имеющуюся в предложении модальность [7, с.183]. Стоит отметить, что это справедливо не для всех функций, и не для всех частиц и случаев употребления.

Исследование функций фразовых частиц разных исторических периодов показало, что исследование фразовых частиц на основе текстов позволяет четче определять их функции, увидеть, как «работают» частицы в тексте, связывают его смысловые части между собой [5]. Однако исследование функций фразовых частиц только на основе художественных произведений имеет свои ограничения. Во-первых, любой художественный текст является отражением языка автора, т.е. это дискурс одного человека, отражающий особенности его речи или речи, которой он подражает. Во-вторых, объем художественных произведений недостаточен, чтобы иметь сравнимый по объему материал для анализа каждой частицы, так как частотность их функций в речи различна. Например, в диалогах будет больше частиц употреблено, чем в повествовании, но чаще тех, что имеют экспрессивные функции. Такое неравномерное распределение может исказить представление о частотности тех или иных функций.

В разрешении представленных выше ограничений может помочь проведение исследования функций фразовых частиц на основе лингвистических корпусов. Для современного китайского языка выделено 10 лингвистических корпусов, к которым имеется открытый доступ [2]. В целом они не имеют существенных различий для исследования функций фразовых частиц, так как имеют схожие лингвистические разметки и поисковые менеджеры.

Использование лингвистического корпуса в исследовании функций фразовых частиц на основе лингвистического корпуса имеет следующие преимущества: большой объем, тексты даны полностью, в ряде случаев можно выбрать конкретные тексты, периоды, материалом является естественный текст, созданный не в учебных целях или для примера, что особенно важно для корпуса разговорной речи. Исследование функций фразовых частиц на основе лингвистического корпуса, также, как и исследование на основе художественного

произведения, проводится на конкретном языковом материале и его синтаксические особенности непременно будут влиять на то, какие частицы используются.

Рассмотрим частотность использования фразовых частиц в разных корпусах, созданных Пекинским университетом языка и культуры (北京语言大学语言智能研究院语料库) [17]. Объем корпуса публицистических текстов составляет 2 млрд. иероглифов, художественной литературы - 3 млрд. иероглифов, разговорного корпуса - 600 млн иероглифов (посты в социальной сети вэйбо, субтитры сериалов и кино).

Таблица 1
Частотность использования фразовых частиц в разных корпусах

Корпус публицистических текстов			Корпус разговорной речи			Корпус художественной литературы		
Фразовая частица	Количество примеров	В процентах	Фразовая частица	Количество примеров	В процентах	Фразовая частица	Количество примеров	В процентах
Всего	2 295 306		Всего	19 200 376		Всего	1 423 306	
了	1607789	70,05%	了	3546149	18,47%	了	786985	55,29%
呢	214186	9,33%	啊	1941774	10,11%	呢	133825	9,4%
吗	122418	5,33%	吧	1067497	5,56%	吧	128926	9,06%
吧	86979	3,79%	吗	1033276	5,38%	吗	126944	8,92%
啊	69273	3,02%	呢	672521	3,5%	啊	72294	5,08%
呀	40344	1,76%	啦	560855	2,92%	呀	41265	2,9%
啦	28941	1,26%	么	465200	2,42%	啦	36899	2,59%
么	14044	0,61%	呀	444921	2,32%	么	15582	1,09%
呵	13097	0,57%	嘛	328582	1,71%	嘛	11952	0,84%
哩	12390	0,54%	呗	58894	0,31%	哩	10445	0,73%

Как мы видим из таблицы 1, несмотря на то, что разговорный корпус меньше в объеме, частицы используются почти на 90% чаще. Также стоит обратить внимание на распределение частотности употребления частиц внутри каждого из корпусов. Если в корпусе художественной литературы и публицистических текстов мы видим преимущественное использование частицы 了 (мы связываем это с тем, что повествовательных предложений в этих корпусах будет больше), а затем шесть частиц, частотность которых постепенно снижается. То в корпусе разговорной речи процент использования каждой из частиц распределен более равномерно. И наиболее частотными частицами в разговорном корпусе становятся частицы, которые чаще используются в выражении экспрессии и паузы (啊, 吧).

В нашем исследовании функций фразовых частиц современного китайского языка (путунхуа) мы анализировали корпус публицистических текстов на основе

текстов ежедневной газеты «人民日报» («Народная ежедневная газета») за 2017–2018 гг. (приблизительно 10 308 газетных страниц, 36 752 639 иероглифов). Так как тексты газеты написаны на нормативном китайском языке (путунхуа), проверены редакторами, оформлены по правилам пунктуации, - все это обусловило выбор именно такого материала для проведения исследования.

Суммарно количество проанализированных примеров по всем частицам составило 36 724 примера. Количество примеров по частицам получилось неравномерным: 啊 (948), 呢 (1947), 吗 (2188), 吧 (884), 的 (11968/814), 了 (18789). Мы связываем преобладание частиц 了, 的 с тем, что характер текста в большей степени повествовательный.

В Таблице 1 и в целом во всем списке фразовых частиц, который автоматически предлагает корпус, не будет указана частица 的. Это связано с технической особенностью при автоматической разметке текста. В нашем исследовании мы столкнулись с тем, что невозможно задать в поисковой строке корпуса параметр «искать иероглиф 的 со значением фразовой частицы», можно только задать поиск по части речи в сочетании с фразовой частицей. В результате мы получили первоначально 52 559 примеров, где встречается и частица 的, и служебное слово 的. Мы разработали алгоритм поиска фразовой частицы по ее позиции в предложении, и это позволило сократить количество примеров до 11 968. И далее уже можно было применять метод экстрадиции примеров с одинаковыми синтаксическими структурами. В случае с частицей 的 оказалось, что в большинстве случаев она будет частью структуры 是...的. После выделения этой структуры осталось лишь 814 примеров для анализа, что было сопоставимо по количеству примеров с другими частицами.

Результатом анализа фразовых частиц на материале лингвистического корпуса стало более подробное описание случаев употребления каждой частицы, выявлены новые устойчивые сочетания в составе синтаксических структур, функции, ранее не описанные в теоретической литературе, уточнены особенности использования фразовых частиц в разных позициях в предложении, в разных типах предложений. Например, для частицы 吗 мы выделили употребление в вопросительных предложениях с вопросительными словами (58 примеров, большая часть из них с вопросительным словом 什么), где фразовая частица усиливает модальность. В анализе частицы 啊 была выделена функция, которая исторически существовала у предшественников современной частицы, но не отмечена в современных справочниках: использование частицы в структурах «существительное—частица—существительное», «глагол — частица — глагол», где глагол или существительное одинаковы. При использовании с глаголами это удвоение глагола; при использовании с существительными это могут быть обращения, рассуждения. В случае фразовой частицы 的 мы подтвердили возможность использования ее в середине предложения (в конце фразового единства), что не указано в современной научной и справочной литературе [7].

Для проведения дальнейших исследований фразовых частиц на материале лингвистического корпуса мы разработали алгоритм анализа, включающий три этапа.

На первом этапе происходит выявление всех примеров с использованием фразовой частицы с помощью встроенного в корпус поискового менеджера. Если количество примеров велико, стоит уже на этом этапе производить поиск в разных типах предложений по-отдельности, сочетая этот поиск с позицией частицы в предложении. Затем примеры необходимо разделить по цели высказывания на повествовательные, вопросительные, побудительные предложения, по интонации на два типа —восклицательные и невосклицательные, по позиции частицы в предложении: в конце предложения или в конце фразового единства, т.е. в середине предложения. Так как лингвистические корпуса не имеют автоматизированного способа выявления вышеописанных типов предложений, необходимо использовать пунктуацию в поисковом запросе.

Второй этап подразумевает первичную обработку примеров, а именно устранение дубликатов, выявление примеров, в которых иероглиф частицы используется в другом значении (например, как словообразовательный суффикс: 酒吧, 书吧).

Третий этап подразумевает анализ примеров. Сначала из массива примеров экстрадируются те, где фразовая частица является частью устойчивой синтаксической структуры, используется с определенными маркерами (наречия, служебные слова и т. д.). Оставшиеся примеры анализируются вручную. Перед началом анализа необходимо разработать кодификатор уже описанных в научной литературе функций. В процессе анализа те примеры, что не могут быть классифицированы в одну из уже имеющихся функций, выделяются отдельно для повторного анализа и поиска схожих грамматических, лексических явлений, анализируется то, какое значение придает частица предложению, как меняется оно, если частицу не использовать.

Исследование функций фразовых частиц китайского языка на материале лингвистических корпусов открывает большие возможности по систематизации всех функций фразовых частиц во множестве контекстов и, вероятно, позволит создать наиболее подробное и точное описание функций на основе больших массивов данных. Лингвистические корпуса позволяют проводить сравнения между функциями частиц разных дискурсов, что позволяет получить более точное, детальное и объективное представление о функциях частицы в разных дискурсах, типах предложений, позициях и сочетаниях. Тем не менее необходимо помнить, об особенностях использования фразовых частиц в разных стилях речи, и что исследование на основе лингвистического корпуса существенно облегчает задачу по анализу, но не решает ее автоматически только лишь за счет объема материала.

Анализ функций фразовых частиц на большом массиве данных не может быть автоматизирован на современном этапе. Однако все большее количество обнаруживаемых закономерностей в построении предложений с частицами в синтаксических структурах дают возможность алгоритмизировать и ускорить обработку примеров.

Литература

1. Горелов В. И. Грамматика китайского языка. Москва: Просвещение, 1982. 279 с.
2. Гурулева Т. Л. Цифровые ресурсы переводчика китайского языка: учебное наглядное пособие. Москва: ВКН, 2022. 158 с.
3. Курдюмов В. А. Морфологический уровень китайского языка как изолирующего топикового. Сущность и содержание позиции частицы // Вестник МГПУ. Сер. Филология. Теория языка. Языковое образование. 2014. № 1(13). С. 37–46.
4. Павлова А. Д., Гурулева Т. Л. Особенности генезиса классических фразовых частиц китайского языка // *Litera*. 2021. № 1. С. 134–143.
5. Павлова А. Д. Развитие научных представлений китайских лингвистов о роли фразовых частиц в китайском языке // Сборник XII Международной научно-практической конференции «Россия — Китай: история и культура». 2019. С. 591–596.
6. Павлова А. Д. Генезис функций фразовых частиц (на материале китайского языка): диссертация ... кандидата филологических наук. Москва, 2021. 274 с. URL: https://vumo.mil.ru/upload/site57/document_file/Pavlova_Avtoreferat.pdf (дата обращения: 10.06.2022).
7. Тань Аошуан. О фразовых модальных частицах (ответ коллеге Юй Сухуа) // Изучение китайского языка. 2003(а). С. 47–52.
8. Шутова Е. А. Синтаксис современного китайского языка. Монография. Москва: Наука, 1991. 392 с.
9. Chao Y. R. *A Grammar of Spoken Chinese*. University of California Press. 1968. 1109 p.
10. Li Boya, *Chinese Final Particles and the Syntax of the Periphery*, 2006. URL: <https://scholarlypublications.universiteitleiden.nl/handle/1887/4393> (дата обращения: 25.07.2019).
11. Li Ch. and Thompson S. *Mandarin Chinese: A Functional Reference Grammar*. University of California Press. 1981. 691 p.
12. Ruey-Jiuan Regina Wu. *A conversation analysis of Mandarin final particles*. John Benjamins Publishing Company. 2004. 260 p.
13. 吕叔湘, 《现代汉语八百词 (增订本)》, 北京: 商务印书馆, 1999. 760 页 [Люй Шусян, 800 слов современного китайского язык. Дополненное издание. Пекин: Изд. Коммерция, 1999. 760 с.]
14. 北京语言大学语言智能研究院语料库 / [Лингвистический корпус китайского языка / Исследовательский институт языкового интеллекта Пекинского университета языка и культуры]. URL: <http://bcc.blcu.edu.cn/> (дата обращения: 10.06.2022).