

ФИЛОЛОГИЯ

DOI: 10.24411/2618-9453-2019-10041

И.В. Самарина

УСТНЫЙ НАРРАТИВНЫЙ ДИСКУРС: КОРПУСНОЕ ИССЛЕДОВАНИЕ ЯЗЫКА ТХО-НЬЫСУАН

Аннотация. В статье описываются предварительные результаты исследования корпуса устных нарративов вьетского языка тхо-ньысуан, записанных во время полевой работы 2018 г. в пров. Тхань-хоа (Вьетнам). Отмечаются различия в наборе критериев для выделения элементарных дискурсивных единиц в процессе сегментации дискурса в языках разной типологии. Рассматриваются особенности функционирования дискурсивных маркеров в языке тхо-ньысуан.

Ключевые слова: устный нарративный дискурс, дискурсивный маркер, изолирующий язык, тональный язык, вьетские языки, язык тхо-ньысуан, народ тхо, корпусный метод, элементарная дискурсивная единица, информационная структура.

Введение

В данной статье на материале изолирующего вьетского языка тхо-ньысуан затрагивается ряд вопросов, связанных с дискурсивной структурой устного нарратива. Как отмечают В.И. Подлеская и А.А. Кибрик, «устная форма языка является основной, фундаментальной и исходной» [Рассказы о сновидениях 2009: 26], а для бесписьменных языков, каковым является тхо-ньысуан, — к тому же и единственной формой существования языка.

В последнее время проблемы дискурса привлекают активное внимание учёных, в то же время исследований, в которых дискурс рассматривается в изолирующих языках, не так много. Отметим работы Чыонг Ван Тиня [Trương Văn Chính 1970], М. Кларка [Clark 1992; 1996], Као Суан Хао [Cao Xuân Hao 1998; 1999], Нгуен Тхи Тхин [Nguyễn Thị Thìn 2002], Буй Тхи Хоанг Ань [Bùi Thị Hoàng Anh 2015], А. Мишо и М. Брюнелля [Michaud, Brunelle 2016: 2048—2070], З.Т. Зо-Хюринвила и Зео Хью Линя [Do-Hurinvillie Danh Thành, Dao Huy Linh 2019], в которых рассматриваются вопросы коммуникативной организации предложения, особенности функционирования дискурсивных маркеров во вьетнамском языке, в том числе в зависимости от их позиции в высказывании, хотя в большинстве работ собственно термин «дискурсивный маркер» не используется.

В процессе изучения особенностей устного дискурса языка тхо-ньысуан методологически мы опирались на корпусный [Рассказы о сновидениях 2009: 27] и сопоставительный методы. Исследование устного нарративного дискурса языка тхо-ньысуан проводилось на полевых материалах, собранных участниками российско-вьетнамской лингвистической экспедиции 2018 г., в том числе вьетнамскими участниками Та Куанг Тунгом (Tạ Quang Tùng), Та Ван Тхонгом (Tạ Văn Thông) и Нгуен Хью Хоанем (Nguyễn Hữu Hoàng), в уезде Ньысуан

пров. Тханьхоа. Настоящее исследование является первым, в котором идиом тхо-ньысуан выступает в качестве языка-объекта. Как показывают наши данные, тхо-ньысуан, скорее всего, является локальным диалектом вьетнамского языка [Самарина 2018: 91], в связи с этим в целях сопоставления и верификации выводов широко привлекались материалы вьетнамского литературного языка.

В качестве компьютерного анализатора речи использовалась программа Speech Analyzer, свободно распространяемая Летним институтом лингвистики (URL: <https://software.sil.org/speech-analyzer/>).

Идиом тхо-ньысуан

Носителями исследуемого идиома являются тхо, проживающие в горном уезде Ньысуан (вьетн. *Như Xuân*) пров. Тханьхоа, который расположен на севере Центрального Вьетнама. Общая численность населения уезда Ньысуан — более 72 тыс. человек. Из этносов представлены кинь, тхай, тхо, мыонг и тай. Как и другой вьетский народ тьыт [Бабаев, Самарина 2018: 14—19], народ тхо объединяет в себе несколько локальных групп — кео, хо, мон, куой, данлай, лиха, поонг (или тай-поонг) и др. [Về vấn đề xác định 1975: 447—451], говорящих на идиомах, которые имеют разную степень близости между собой и языками вьет-мыонгской подгруппы — вьетнамским и мыонг. Вьетнамский исследователь Нгуен Хыу Хоань показал, что идиом кео является локальным диалектом вьетнамского языка; идиомы хо и мон относятся к диалектам языка мыонг; идиомы тхо (уезды Ламла и Ньысуан пров. Тханьхоа), куой и поонг-данлай следует считать самостоятельными языками [Нгуен Хыу Хоань 2018: 87]. Однако, по данным наших полевых материалов 2017 и 2018 г., языки поонг и данлай являются самостоятельными, хотя и близкородственными — 84 % общей лексики по 100-словному списку М. Сводеша, а идиом тхо уезда Ньысуан следует считать локальным диалектом вьетнамского языка — 95 % общей лексики по 100-словному списку М. Сводеша [Самарина 2018: 91]. Поскольку известны ещё и другие идиомы тхо (например, тхо-шонгкон, являющийся диалектом языка мыонг) [Mạc Đường 1997: 88—90], то при их рассмотрении, во избежание неоднозначности в установлении того, о каком именно идиоме тхо идёт речь, в их названиях мы будем использовать уточняющие компоненты. Так, в статье речь идёт об идиоме тхо-ньысуан.

Тхо-ньысуан является моносиллабическим изолирующим языком. Это означает, что простое слово в нём, как правило, материально совпадает со слогом и морфемой, на роль слов в предложении указывает порядок слов и служебные элементы. Обычный порядок слов в клаузе — SOV, определяющие или уточняющие компоненты располагаются в постпозиции к главному, для высказывания типична топико-комментная структура.

Сегментация дискурса

Первостепенным вопросом при рассмотрении дискурса является сегментация речевого потока на элементарные дискурсивные единицы (ЭДЕ), т. е. «на кванты, обладающие просодическим, семантическим и синтаксическим единством» [Рассказы о сновидениях 2009: 47].

В качестве главного просодического (интонационного) критерия обычно выбирается частота основного тона [Рассказы о сновидениях 2009: 57]. Однако следует отметить, что в тональных языках частота основного тона — это характеристика базовой лексико-грамматической единицы (фонологического слова), которая несёт на себе лексический тон и в ка-

честве которой обычно выступает слог или полутораслог (sesquisyllabic unit), поэтому в тональных языках критерий «единый интонационный контур» для вычленения ЭДЕ не применим.

Частое совпадение ЭДЕ с клаузой позволяет при сегментации дискурса использовать в качестве критерия наличие в ЭДЕ финитной формы глагола, однако в изолирующих языках подобный критерий также не работает.

При этом среди компонентов дискурса всё же имеются пограничные сигналы для ЭДЕ, встречающиеся в языках любой типологии, — это дискурсивные маркеры (незнаменательные слова или словосочетания), а также паузы.

Паузы являются обязательной частью дискурса. Это обусловлено физиологическими особенностями порождения речи, а именно необходимостью периодических вдохов. Пауза может использоваться также как вспомогательное средство в случае хезитации. «Членение речи отчасти объясняется физиологической потребностью периодически делать вдохи, отчасти спланировать следующий квант дискурса» [Рассказы о сновидениях 2009: 29].

Отсутствие в изолирующих языках формальных средств связи между клаузами (союзов, финитных форм глагола и пр.) затрудняет объединение клауз в предложения, и дискурс изолирующего языка предстаёт как последовательность ЭДЕ, связанных между собой теми или иными дискурсивными значениями, что ещё раз подчёркивает значимость для изолирующих языков коммуникативной организации, информационной структуры (information structure) [The Oxford Handbook 2016] предложения.

Дискурсивные маркеры

Под дискурсивными маркерами вслед за А.А. Кибриком мы будем понимать «слова или словосочетания, кодирующие значения, отличные от пропозиционального содержания (или от истинностной оценки)» [Кибрик 2009: 3]. Имеются различные подходы к рассмотрению дискурсивных маркеров (см., например, небольшой обзор Е.М. Мурашковой [Мурашкова 2014]). Дискурсивные маркеры обладают различными функциями, в том числе служат как средства организации дискурса [Агранат 2009: 65] или связности текста [Мурашкова 2014: 53], для эмфатического выделения какого-либо из членов предложения [Агранат 2009: там же], являются средствами хезитации, используемыми для заполнения временного промежутка, когда человек испытывает трудности в дальнейшей реализации речевого замысла, когда ему требуется пауза для обдумывания речевой стратегии или подбора лексики.

Принадлежность лексической единицы к классу дискурсивных маркеров определяется главным образом на основании функциональных критериев. При употреблении лексических единиц в качестве дискурсивных маркеров происходит их полная или частичная десемантизация, при этом в других контекстах они продолжают использоваться в качестве значимых единиц. Так, в примере 1 компоненты *keu¹* ‘звать[ся]’ (вьетн. *kêu* ‘кричать’, ‘издавать звуки’) и *la²* ‘быть’ употреблены в своих лексических значениях, в то время как в примере 2, компоненты дискурсивных комплексов *keu¹ la²* и *vɔi⁴ la²* (букв. ‘звать[ся]’ + ‘быть’), так сказать, подверглись десемантизации (правила оформления примеров см. ниже в разделе «О транскрипции и структуре примеров»).

- | | | | | | |
|----|------------------------|-----------------------|------------------------|-----------------------|-----|
| 1) | <i>lan²</i> | <i>ni¹</i> | <i>keu¹</i> | <i>la²</i> | ... |
| | làng | này | gọi | là | |
| | деревня | этот | называть[ся] | быть; COP | |

Эта деревня называется...

2.1) $v\text{oi}^4\text{-}la^2$ ti^2 nu^1 $\eta a^{\check{y}2}$ ce^1 la^2
 gọi là đi như ngày xưa là
 так сказать; DM идти как день прежние_времена сор; DM

2.2) $keu^1\text{-}la^2$ ||
 gọi là
 так сказать; DM

2.3) ti^2 $t^h a\check{m}^2$ $t^h a\eta^3$ ||...
 đi thăm tháng
 идти посещать месяц

Так сказать, идти как в старину... так сказать, совершать [еже]месячное посещение... (речь идёт об обряде ежемесячного посещения дома невесты и её родителей родственниками жениха в течение года, предшествующего свадьбе. — И.С.).

Дискурсивным маркерам во вьетнамском языке посвящена диссертация Буй Тхи Хоанг Ань, в которой рассмотрены как составные дискурсивные маркеры (например, *tuỳ nhiên* ‘однако’, ‘тем не менее’, *tuỳ thé* и *tuỳ vậу* ‘несмотря на’, *dù sao* ‘всё равно’, ‘всё же’, *cũng vẫn* ‘всё равно’, ‘всё-таки’ и др.), так и проанализированы дискурсивные значения отдельных компонентов, в том числе *thì* ‘тогда’, *mới* ‘только тогда’, ‘только что’, *mà* ‘но’, ‘однако’, ‘который’, *vẫn* ‘всё-таки’, ‘тем не менее’ и т. д. [Bui Thi Hoang Anh 2015].

Анализ корпуса текстов на языке тхо-ньысуан показал, что наборы дискурсивных маркеров во вьетнамском языке и идиоме тхо-ньысуан имеют много общих элементов, ср.: тх. $t^h e^3$ ~ вьетн. *thé* ‘так’; тх. ci^2 ~ вьетн. *thì* ‘тогда’; тх. ma^2 ~ вьетн. *mà* ‘но’, ‘однако’; тх. la^2 ~ вьетн. *là* ‘быть’ (*связка*); тх. $m\check{x}^3$ ~ вьетн. *mới* ‘только тогда’, ‘только что’; тх. $t^h oi^1$ ~ вьетн. *thôi* ‘лишь’; тх. roi^2 ~ вьетн. *rôi* ‘потом’; тх. $\beta\check{x}^4$ ~ вьетн. *vậу* ‘так’; тх. $\beta\check{m}^4$ ~ вьетн. *vẫn* ‘всё-таки’, ‘тем не менее’; тх. $ku\eta^4$ ~ вьетн. *cũng* ‘также’; тх. na^3 ~ вьетн. *đây* ‘то’; тх. ni^1 ~ вьетн. *đây* ‘это’ (во всех этих случаях приведены лексические, а не дискурсивные значения компонентов). Наблюдается схожесть в развитии дискурсивных значений и в особенностях функционирования этих элементов. Так, Буй Тхи Хоанг Ань, рассматривая дискурсивные маркеры *thì* ‘тогда’, *mà* ‘однако’, *là* ‘быть’, *mới* ‘только тогда’, ‘только что’, отмечает их рематический характер [Bui Thi Hoang Anh 2015: 29]. В своё время Као Суан Хао также писал, что компонент *thì* ‘тогда’ указывает на границу между темой и ремой, и сделал при этом интересное замечание о том, что в фонетической синтагме этот компонент относится к реме, поскольку произносится с малой интенсивностью [Сао Хуân Hào 1999: 125]. В языке тхо-ньысуан мы обнаруживаем схожее явление. Ниже на рис. 1 изображено окно звукового редактора Speech Analyze 3.1, в котором показана осциллограмма, тональный контур и интенсивность произнесения высказывания (пример 3), в составе которого имеется рематический дискурсивный маркер ci^2 ‘тогда’, разделяющий это высказывание на тему ($la^2 mo^1 ma^2 \beta it^3$ ‘в каком доме мало’) и рему ($du\check{r}\check{x}^4 han^1 la^3$ ‘получаться два листа’), при этом дискурсивный маркер ci^2 (на рис. 1 он выделен курсорами) произносится с крайне низкой интенсивностью, присоединяясь, по фонетическим законам организации синтагмы в мон-кхмерских языках (= сильная позиция находится в конце синтагмы), к реме. Обратим внимание на то, что в примере 3 имеется еще один тема-рематический дискурсивный маркер ma^2 , который в теме $la^2 mo^1 ma^2 \beta it^3$ ‘в каком доме мало’ отделяет подтему $la^2 mo^1$ ‘дом какой’ от подремы βit^3 ‘мало’.

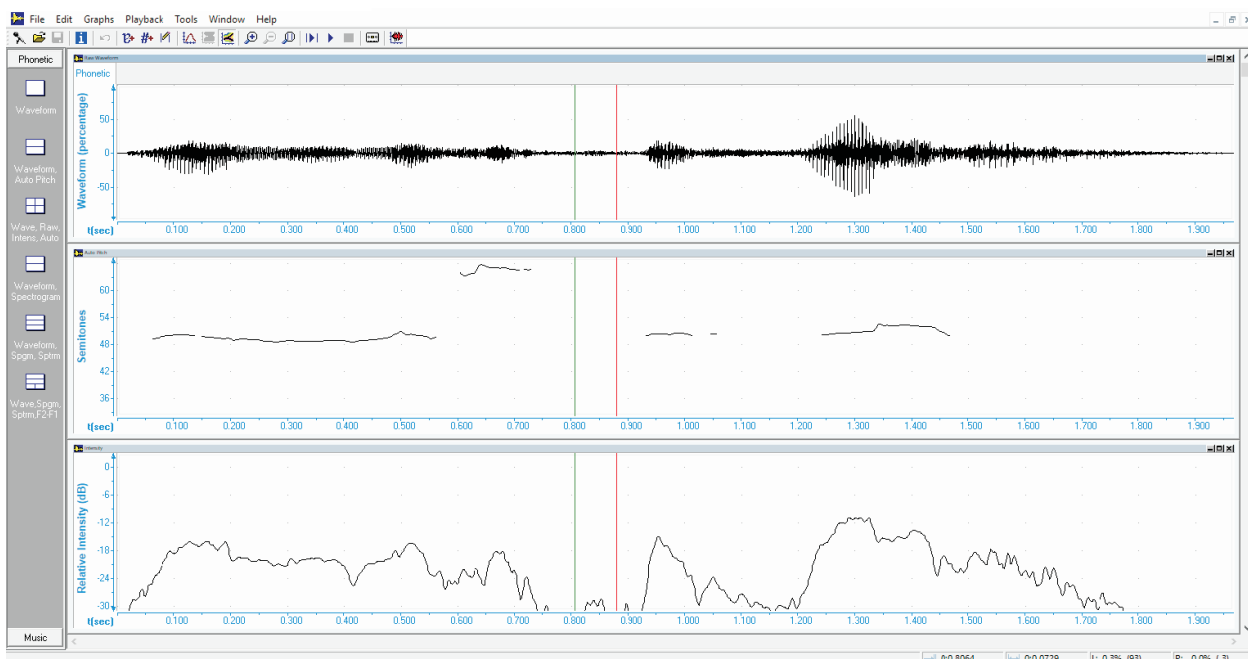


Рис. 1. Дискурсивный маркер *ci¹* ‘тогда’ в составе ЭДЕ

3) *la²* *to¹* *ta²* *lit³* *ci¹* *dʃɤk* *ha : n¹* *la³*
 дом какой DM мало DM получать[ся] два лист

Какие дома бедные, то готовят два листа (= порции) (о подготовке свадебного угощения, которое подаётся на листьях растения. — И.С.).

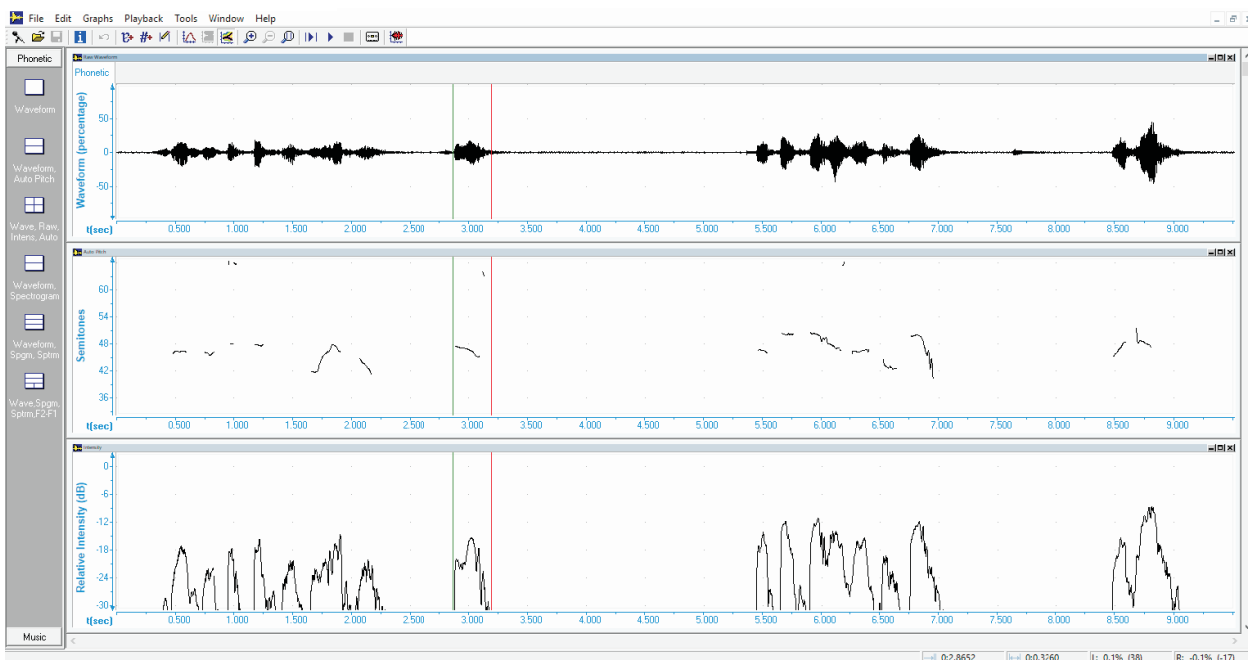


Рис. 2. Дискурсивный маркер *ci¹* ‘тогда’ как отдельная ЭДЕ

Один и тот же маркер может как входить в состав ЭДЕ, так и самостоятельно её образовывать. На рис. 2 в окне редактора Speech Analyzer 3.1 показаны акустические параметры высказывания (транскрипцию и перевод см. в примере 4), состоящего из четырёх ЭДЕ, разделённых паузами, и одной из этих ЭДЕ является дискурсивным маркером *ci¹* ‘тогда’ (на рисунке выделен курсорами).

4.1) $\epsilon\epsilon u^1$ $f\alpha\eta^1-tuk^4$ $t\check{\alpha}^4-ku\alpha n^3$ $ku\alpha^4$ $\eta\alpha\alpha i^2$ t^ho^4 |
 theo phong_tục tập_quán của người Thỏ
 следовать обычай привычка POSS человек тхо

4.2) ϵi^1 ||
 thì
 тогда;DM

4.3) $t^hu^4-tuk^4$ $t\alpha^2$ $\eta\alpha i^2$ $\epsilon\epsilon^1$ $te\eta^3$ cu^2 ||
 thủ_tục từ ngày xưa đến giờ
 процедура ABL день давний приходит время

4.4) $\eta\alpha i^2$ $tlu\alpha k^3$ ||
 ngày trước
 день перёд

В соответствии с традициями и обычаями у тхо [свадебная] процедура с древних времен до последнего времени [вот какая была]

Дискурсивные маркеры в языке тхо-ньысуан бывают больше, чем слово, например дискурсивный маркер $n\alpha i^3 \beta\epsilon n^2$ (букв. ‘говорить’ + ‘возвращаться’) ‘что касается’, ‘говорить о’, который употребляется в начале ЭДЕ и указывает на левую границу темы; см. пример 5:

5.1) $n\alpha i^3$ $\beta\epsilon n^2$ $f\alpha\eta^1-tuk^4$ $t\check{\alpha}^4-quan^3$ $ku\alpha^4$ $\eta\alpha\alpha i^2$ t^ho^4
 nói về phong_tục tập_quán của người Thỏ
 говорить возвращаться обычай привычка POSS человек тхо

5.2) la^2 $k\epsilon^4$ $ti\eta^3$ $du\alpha n^2-ket^3$ $\eta\alpha i^2$ $\epsilon\epsilon^1$ la^2 $ka\alpha i^1$
 là cái tình đoàn_kết ngày xưa là cao
 COP; DM CL чувство солидарность день давний COP; DM высокий

[Если] говорить о привычках и обычаях тхо, то в прежние времена чувство солидарности было очень высоким.

В приведённом примере на правую границу темы указывает дискурсивный маркер la^2 , который относится к другой ЭДЕ и стоит в её начале. Второй дискурсивный маркер la^2 в ЭДЕ в примере 5.2 указывает на границу между подтемой и подремой в этой клаузе.

Самым частотным дискурсивным маркером в языке тхо-ньысуан является la^2 ‘быть’. С его помощью указывается на а) тождественность в некотором аспекте фрагментов дискурса; б) то, что последующий фрагмент дискурса поясняет, уточняет предшествующий или просто вводит следующий фрагмент дискурса. Как и дискурсивные маркеры ϵi^2 ‘тогда’, ma^2 ‘но, однако’, $m\alpha i^3$ ‘только что’, ‘только тогда’ маркер la^2 ‘быть’ может служить тема-рематической границей между частями высказывания. Кроме того, la^2 используется также в качестве хезитационного средства, когда говорящий испытывает те или иные сложности с разворачиванием повествования. В этом случае он фонетически удлиняется (см. примеры 6.1—6.3).

- 6.1) *la-a-a²* ||
là
COP; DM
- 6.2) *lu¹ ηαι̋² ɛe¹ ti² kɔ³ ||*
 như ngày xưa đi có
 COMP день давние_времена идти иметь[ся]
- 6.3) *mo⁴ kɛ⁴ ʔoη¹ mɔ¹ | ...*
 một cái ông mo
 один CLF дед шаман

Сватовство... как в прежние времена, [когда] шли, был шаман...

В случае хезитационного употребления дискурсионных маркеров, т. е. их употребления как вспомогательного средства, когда говорящий испытывает различного рода затруднения в процессе порождения последующего фрагмента дискурса и ему требуется время на подбор лексических единиц, обдумывание стратегии описания ситуации, исправление речевых сбоев и пр., нередки случаи, когда подряд следует несколько дискурсионных маркеров. При этом речь идёт не о составных дискурсионных маркерах, состоящих из простых (одно-сложных) маркеров, например *kuη⁴ βη̋⁴* (букв. ‘также’ + ‘всё-таки’) ‘всё также’, не об употреблении последовательности дискурсивных маркеров, связь между которыми если не идиоматична, то по крайней мере их употребление можно считать связанным или логически обусловленным, например *na³ la²* (букв. ‘то[т]’ + ‘быть’) ‘то’, ‘то есть’, *ci² mη̋³* (‘букв. ‘тогда’ + ‘только тогда’) ‘то только тогда’, а именно об употреблении последовательности дискурсионных маркеров, выбор которых скорее обусловлен хезитацией (см. примеры 7 и 8).

- ɔ-ɔ-* ||
- 7.1) *ɔ¹ la² keű¹ na³ la²*
 σ-σ-σ là kêu đấу là
DM COP;DM называть[ся];DM PRT;DM COP;DM
- 7.2) *ɛεű¹ kɛ⁴ fɔη¹-tuk⁴ ηαι̋² ɛe² la²*
 theo cái phong_tục ngày xưa là
 следовать CLF обычай день прежние_времена **COP;DM**

... следуя обычаям прежних дней...

Как видно из примеров 8.1—8.2, говорящий начинает формировать квант дискурса, однако не может сразу выбрать нужный вариант, и, наконец, после абсолютной паузы, явно имеющей и хезитационное значение, а не только физиологическое, следует каноническая ЭДЕ (8.2).

- 8.1) *cu³ k^hoη¹ kɔ³ βη̋³-de² ci¹ vɔ̋⁴-la² na³ ||*
 chú không có vấn_đề chi gọi_là đố
 a;conj NEG иметь[ся] проблема что так_сказать;DM тот;DM

| | | | | | |
|------|--------------------------|-----------------------|------------------------|------------------------------------|--|
| 8.2) | <i>cuj m⁴</i> | <i>ti²</i> | <i>han¹</i> | <i>k x₂³</i> | |
| | chuyện | đi | hỏi | vợ | |
| | дело | идти | спрашивать | жена | |

И нет никаких проблем, скажем так идти свататься.

При исследовании дискурсивных маркеров следует обращать внимание на позицию, в которой употреблён тот или иной маркер (в начале, середине или конце ЭДЕ), входит в состав ЭДЕ или образует как самостоятельную. От этого может зависеть значение дискурсивного маркера и особенности его функционирования.

Выводы

Поскольку язык тхо-ньёсуан можно считать диалектом вьетнамского языка, то, естественным образом, наборы дискурсивных маркеров и их функционирование в обоих языках во многом совпадают.

Так как основной просодический критерий сегментации ЭДЕ — единый интонационный контур, определяемый частотой основного тона, — не является значимым для тональных языков, то следует более глубоко рассмотреть вопрос о релевантных просодических характеристиках, которые могут учитываться при вычленении ЭДЕ. Это может быть интенсивность произнесения, темп речи, тональные изменения.

О транскрипции и структуре примеров

Примеры из вьетнамского языка приводятся в орфографии. Все примеры из языка тхо-ньёсуан записаны в статье знаками Международного фонетического алфавита (IPA). Помимо знаков для сегментных единиц, используются два знака для супrasegmentных единиц: знак | — для обозначения краткой паузы и знак || — для обозначения длительной паузы. Тоны в языке тхо-ньёсуан обозначаются надстрочными арабскими цифрами. Другие просодические характеристики, например интенсивность, не отмечаются.

Для каждого примера фрагмента дискурса, помимо транскрипции, ниже приводится пословный перевод на вьетнамский литературный язык и глоссирование на русском языке в соответствии с Лейпцигской системой правил глоссирования (URL: <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>), а также литературный перевод на русский язык. Используемые сокращения см. ниже.

Номера примеров приводятся за круглой скобкой. Если пример содержит более одной ЭДЕ, то каждая из них записывается с новой строки и нумеруется отдельно: после номера примера ставится точка, за которой указывается порядковый номер ЭДЕ. Так, нотация 4.3 означает, что это третья ЭДЕ примера 4.

Сокращения

- вьетн. — вьетнамский
- тх. — тхо-ньёсуан
- ЭДЕ — элементарная дискурсивная единица
- ABL — аблатив
- CLF — классификатор

- COM — комитатив
 CONJ — союз
 COP — связка
 DM — дискурсивный маркер
 F — женский пол
 ITER — итератив
 M — мужской пол
 NEG — отрицание
 ORD — показатель порядкового числительного
 POSS — посессив
 PRT — частица
 | — краткая абсолютная пауза
 || — длительная абсолютная пауза.

Список литературы

1. Агранат Т.Б. Дискурсивные маркеры в водском языке // *Вопросы языкознания*, 2009, № 6. С. 65—75.
2. Бабаев К.В., Самарина И.В. Язык май. Материалы российско-вьетнамской лингвистической экспедиции. Выпуск 5. М.: ЯСК, 2018. 576 с.
3. Кибрик А.А. Лекция «Дискурсивные маркеры» 2009 г. URL: <http://www.philol.msu.ru/~otipl/new/main/courses/discourse/New/14-discourse-markers.pps> (дата обращения: 27.10.2019).
4. Мурашкова Е.М. Подходы к исследованию дискурсивных маркеров // *Вестник МГОУ. Серия: Лингвистика*, 2014, № 5. С. 53—59.
5. Нгуен Хыу Хоань. Проблемы языковой принадлежности народа тхо // *Малочисленные языки и народы Центрального Вьетнама*. Материалы к сопоставительному изучению вьетских языков провинций Нгеан и Тханьхоа. М.: Тезаурус, 2018. С. 73—86.
6. Подлеская В.И., Кибрик А.А. Дискурсивные маркеры в структуре устного рассказа: опыт корпусного исследования // *Компьютерная лингвистика и интеллектуальные технологии*. М.: РГГУ, 2009. С. 390—395.
7. Рассказы о сновидениях: корпусное исследование устного русского дискурса / под ред. А.А. Кибрика, В.И. Подлеской. М.: ЯСК, 2009. 736 с.
8. Самарина И.В. 100-словный список Сводеша для языков данлай, поонг и тхо-ньысуан // *Малочисленные языки и народы Центрального Вьетнама*. Материалы к сопоставительному изучению вьетских языков провинций Нгеан и Тханьхоа. М.: Тезаурус, 2018. С. 87—91.
9. Bui Thi Hoang Anh. Etude des marqueurs discursifs du vietnamien dans une perspective comparative avec des marqueurs discursifs du français / Thèse [Буй Тхи Хоанг Ань. Изучение дискурсивных маркеров вьетнамского языка в сравнении с дискурсивными маркерами французского/ Диссертация]. Université Paris 7, Sorbonne Paris Cité; UFR Linguistique 2015. URL: <https://tel.archives-ouvertes.fr/tel-01536404/document> (дата обращения: 23.10.2019).
10. Cao Xuân Hạo. Về ý nghĩa ‘thì’ và ‘thể’ trong tiếng Việt // *Ngôn ngữ* [Као Суан Хао. О значении thì и thể во вьетнамском языке // *Язык*], 1998, Số 5. Tr. 1—31.
11. Cao Xuân Hạo. Tiếng Việt sơ thảo ngữ pháp chức năng [Као Суан Хао. Очерк функциональной грамматики вьетнамского языка]. Hà Nội: Nxb Giáo dục, 1999. 495 tr.
12. Clark M. Conjunction as copula in Vietnamese // *Mon-Khmer Studies*, 1996, Vol. 26. P. 319—331.
13. Clark M. Conjunction as topicalizer in Vietnamese // *Mon-Khmer Studies*, 1992, Vol. 20. P. 91—109.

14. Do-Hurinville Danh Thành, Dao Huy Linh. The Vietnamese polyfunctional marker mà as a generalized linker: a multilevel approach // *Journal of the Southeast Asian Linguistics Society*, Vol. 12.2.2019. P. 58—71.

15. Mạc Đường. Dân tộc học và vấn đề xác định thành phần dân tộc [Мак Дьонг. Этнография и проблема определения состава народов]. Hà Nội: Nxb Khoa học Xã hội, 1997.

16. Michaud A., Brunelle M. Information Structure in Asia: Yongning Na (Sino-Tibetan) and Vietnamese (Austroasiatic) // *The Oxford Handbook of Information Structure*. Oxford University Press, 2016. P. 774—789.

17. Nguyễn Thị Thìn. Các từ THI, MA, NHƯNG ở đầu câu trong chức năng liên kết nghĩa học // *Ngôn ngữ* [Нгуен Тхи Тхин. Слова thì, mà и nhưng в начале предложения в функции семантических связей // *Язык*], 2002, Số 5. Tr. 31—37.

18. *The Oxford Handbook of Information Structure*. Oxford University Press, 2016. 1186 p.

19. Trương Văn Chỉnh. Structure de la langue vietnamienne / *Publications du Centre Universitaire des Langues Orientales vivantes* [Чьюнг Ван Тинь. Структура вьетнамского языка // *Публикации Университетского центра живых восточных языков*], 6è série, Tome X, 1970. 478 p.

20. Về vấn đề xác định thành phần các dân tộc thiểu số ở miền Bắc Việt Nam [О проблеме определения состава малочисленных народов севера Вьетнама]. Hà Nội: Nxb Khoa học Xã hội, 1975.

Для цитирования: Самарина И.В. Устный нарративный дискурс: корпусное исследование языка тхо-ньысуан // *Вьетнамские исследования*, сер. 2, 2019, № 4. С. 72—83.

Автор:

Самарина Ирина Владимировна, научный сотрудник Института языкознания РАН, старший преподаватель Института классического Востока и античности НИУ ВШЭ. E-mail: ira_samarina@hotmail.com

Продвижение статьи:

Дата поступления: 30.11.2019

Дата поступления в переработанном виде: 05.12.2019

Принята к печати: 11.12.2019

I.V. Samarina

ORAL NARRATIVE DISCOURSE: CORPUS STUDY OF THO (NHU XUAN) LANGUAGE

Abstract. The article describes the preliminary results in corpus study of oral narratives in the Tho (Nhu Xuan) language. The narratives were recorded during the fieldwork 2018 in Thanh Hoa province (Vietnam). Differences in the set of criteria for discourse segmentation in languages of different typologies are noted. The functioning of discursive markers in the Tho (Nhu Xuan) language are considered.

Keywords: oral narrative discours, discursive marker, hesitation, isolating language, tonal language, Vietic languages, Tho (Nhu Xuan) language, Tho ethnic minority, corpus method, elementary discursive unit, information structure.

References

1. Agranat, T.B. (2009). Diskursivnye markery v vodskom yazyke [Discursive markers in Votian language]. *Voprosy yazykoznanija* [Linguistic Issues], 2009. № 6. S. 65—75.
2. Babaev, K.V., Samarina, I.V. (2018). Yazyk maj. Materialy rossijsko-vietnamskoi lingvisticheskoi ekspedicii, vyp.5 [May language. Materials of the Russian-Vietnamese linguistic expedition. Issue 5]. M.: YASK, 2018. 576 s.
3. Bui Thi Hoang Anh. Etude des marqueurs discursifs du vietnamien dans une perspective comparative avec des marqueurs discursifs du français / Thèse [Study of discursive markers of Vietnamese in a comparative perspective with discursive markers of French / Thesis]. Université Paris 7, Sorbonne Paris Cité; UFR Linguistique 2015. Retrieved on October 23, 2019 from URL: <https://tel.archives-ouvertes.fr/tel-01536404/document>.
4. Cao Xuân Hạo (1998).. Về ý nghĩa ‘thì’ và ‘thể’ trong tiếng Việt [About the meanings of ‘thì’ and ‘thể’ in Vietnamese]. *Ngôn ngữ*, 1998, Số 5. Tr. 1—31.
5. Cao Xuân Hạo (1999). Tiếng Việt sơ thảo ngữ pháp chức năng [The outline of the Vietnamese functional grammar]. Hà Nội: Nxb Giáo dục, 1999. 495 tr.
6. Clark, M. (1996). Conjunction as copula in Vietnamese. *Mon-Khmer Studies*, 1996, Vol. 26. P. 319—331.
7. Clark, M. (1992). Conjunction as topicalizer in Vietnamese. *Mon-Khmer Studies*, 1992, Vol. 20. P. 91—109.
8. Do-Hurinville Danh Thành, Dao Huy Linh (2019). The Vietnamese polyfunctional marker mà as a generalized linker: a multilevel approach. *Journal of the Southeast Asian Linguistics Society*, Vol. 12.2. 2019. P. 58—71.
9. Kibrik, A.A. (2009). Lekciya «Diskursivnye markery» 2009 g. [Lecture “Discursive Markers” 2009]. Retrieved on April 27, 2019 from URL: <http://www.philol.msu.ru/~otipl/new/main/courses/discourse/New/14-discourse-markers.pps>.
10. Mạc Đường (1997). Dân tộc học và vấn đề xác định thành phần dân tộc [Ethnography and ethnicity identification issues]. Hà Nội: Nxb Khoa học Xã hội, 1997.
11. Michaud, A., Brunelle, M. (2016). Information Structure in Asia: Yongning Na (Sino-Tibetan) and Vietnamese (Austroasiatic). *The Oxford Handbook of Information Structure*. Oxford University Press, 2016. P. 774—789.

12. Murashkovskaya, E.M. (2014). Podhody k issledovaniyu diskursivnykh markerov [Approaches to the study of discursive markers]. *Vestnik MGOU. Seriya: Lingvistika* [Bulletin of MSRU. Series of Linguistics], 2014, № 5. S. 53—59.
13. Nguyen Huu Hoanh (2018). Problemy yazykovej prinadlezhnosti naroda tho [Problems of the language affiliation of the Tho people]. *Malochislennyye yazyki i narody Centralnogo Vietnama*. Materialy k sopostavitel'nomu izucheniyu vietskiykh yazykov provinciy Ngean i Thanhhoa. M.: Tezaurus, 2018. S. 73—86.
14. Nguyễn Thị Thìn (2002). Các từ THI, MÀ, NHƯNG ở đầu câu trong chức năng liên kết nghĩa học [The words thì, mà, and nhưng at the beginning of a sentence in semantic connective function]. *Ngôn ngữ* [Language], 2002, Số 5. Tr. 31—37.
15. Podlesskaya, V.I., Kibrik, A.A. (2009) Diskursivnye markery v strukture ustnogo rasskaza: opyt korpusnogo issledovaniya [Discursive markers in the structure of oral narratives: the case study of corpus research]. *Kompyuternaya lingvistika i intellektualnye tekhnologii*. M.: RGGU, 2009. S. 390—395.
16. Rasskazy o snovideniyah: Korpusnoe issledovanie ustnogo russkogo diskursa [Dream Stories: Case Study of Oral Russian Discourse] / Pod red. A.A. Kibrika, V.I. Podlesskoj. M.: YASK, 2009. 736 s.
17. Samarina, I.V. (2018). 100-slovnny spisok Svodesha dlya yazykov danlay, poong i tho-nysuan [100-words Swadesh list for Dan Lai, Poong, and Tho (Nhu Xuan) languages]. *Malochislennyye yazyki i narody Centralnogo Vietnama*. Materialy k sopostavitel'nomu izucheniyu vietskiykh yazykov provinciy Ngean i Thanhhoa. M.: Tezaurus, 2018. S. 87—91.
18. The Oxford Handbook of Information Structure. Oxford University Press, 2016. 1186 p.
19. Trương Văn Chính (1970). Structure de la langue vietnamienne [Vietnamese language structure]. *Publications du Centre Universitaire des Langues Orientales vivantes*, 6^e série, tome X. 1970. 478 p.
20. Về vấn đề xác định thành phần các dân tộc thiểu số ở miền Bắc Việt Nam [Regarding the identification of ethnic minorities in Northern Vietnam]. Hà Nội: Nxb Khoa học Xã hội, 1975.

For citation: Samarina I.V. (2019). Oral narrative discourse: corpus study of tho (nhu xuan) language. *Russian Journal of Vietnamese Studies*, series 2, № 4. P. 72—83.

Author:

Samarina Irina V., Researcher, Institute of Linguistics RAS, Senior Lecturer, Institute for Oriental and Classical Studies, National Research University “Higher School of Economics”.
Email: ira_samarina@hotmail.com

Article history:

Received: November 30, 2019

Received in revised form: December 05, 2019

Accepted: December 11, 2019