

Методика корпусных исследований паузации на примере изучения паузации в японском языке в контексте послелогов и топики¹

Pauses after postpositions and topical particle *wa* in Japanese: a corpus study

Комарова А. Д. (komarovochka@gmail.com)

Российский государственный гуманитарный университет

В данной работе описывается паузация в японском языке до и после первичных послелогов и топикиальной частицы. Рассматриваются такие параметры, как вероятность, длина и заполнение данных пауз, а также случаи, когда они оказываются на разных типах синтаксических границ.

В данной работе рассматривается паузация в устном японском дискурсе в позициях до и после приемных послелогов и топикиализующей частицы.

Материалом для исследования служил корпус из 45 тестов — «рассказов по картинкам». Всего от двенадцати говорящих было записано по четыре текста с рассказами и пересказами несложных сюжетов по наборам картинок, однако в связи с плохим качеством записи 3 текста не использовались.

Для каждого из тестов было посчитано абсолютное число вхождений того или иного послелога (частицы), количество пауз после данного показателя, а также случаи, когда в непосредственной близости встречались речевые сбои — заполненные паузы или коррекции.

Паузация представляет интерес как одно из явлений, присущих устному дискурсу. В устной речи паузы обладают рядом функций, например, дают говорящему возможность сделать вдох. В данной работе принят когнитивный подход, в соответствии с которым позиции, в которых возникают паузы, обусловлены, в частности, синтаксическими критериями.

Так, в [Chafe 1994] было показано, что интонационные единицы (т. е., условно говоря, фрагменты от паузы до паузы) в 60% случаев соответствуют клаузам. Аналогичные выводы для русского языка были сделаны в [Кривнова, Чардин 1999]. А в [Кибрик, Подлесская 2009] также указываются такие характерные позиции «планирования», т. е. синтаксические границы, на которых говорящий делает

паузу, позволяющую ему спланировать новый отрезок, как границы предложений.

Данные японского языка представлены в [Iwasaki 1993] и [Sakura, Fuji 2006], причем в первой работе утверждается, что интонационные единицы в японском языке мельче, чем в английском, а во второй показывается, что, как правило, интонационные единицы, меньшие, чем клаузы — это топикиальные группы.

Паузация в японском языке на границах предложений и клауз, входящих в сложные предложения, рассматривалась в [Комарова 2008]. Результаты этой работы показали, что паузы на границах предложений в устном японском дискурсе возникают в 94% случаев и часто (в 23% случаев) являются заполненными, а паузы на границах клауз возникают приблизительно в 60% случаев.

Паузы после топикиализующей частицы и послелогов относятся, как правило, к паузам внутри клауз, но в некоторых случаях могут совпадать с паузами на границах клауз и предложений.

1. Паузы после топикиализующей частицы *wa*

Частица *wa* маркирует топик, т. е. ту именную группу, о которой в предложении что-то сообщается, тему высказывания (см. [Алпатов, Аркадьев, Подлесская 2008] часть 2, стр.36–55). Топикиализующая частица *wa* встретилась в корпусе всего 205 раз.

¹ Работа выполнена при финансовой поддержке РГНФ, грант №09-04-00106а

В 72 случаях после данного показателя были незаполненные паузы. Еще в 27 случаях данные паузы были заполненными или смешанными (т. е. содержали в себе и лексически или долексически заполненный фрагмент, и незаполненный). Таким образом, всего паузы после топиализующей частицы *wa* встретились в 48% случаев (99 раз), причем в 27% случаев эти паузы были заполненными (здесь и в дальнейшем под заполненными паузами будут пониматься и заполненные, и смешанные, т. к. в данном случае принципиально именно наличие «филлера» — заполнителя).

Что касается длины пауз после *wa*, то для незаполненных пауз среднее статистическое составило 0,5 сек, медиана — 0,3 сек, мода — 0,1 сек. Для смешанных пауз эти показатели сильно отличались и составили, соответственно, 1,2 сек, 1,3 сек и 1,4 сек. В данном случае интересно сравнить показатели для незаполненных и смешанных пауз. Сопоставление показывает, что длина стандартной незаполненной паузы значительно (почти на 1 сек.) меньше, чем длина паузы с заполнением. Это показывает, что наличие заполнителя связано с когнитивными затруднениями говорящего и удлиняет паузу. Если рассматривать все (незаполненные и смешанные) паузы вместе, то общие показатели составляют: среднее статистическое — 0,7 сек, медиана — 0,4 сек и мода — 0,1 сек.

С синтактико-семантической точки зрения позиции *wa* в тексте были не равноценными. Так, наиболее характерно употребление топики в начале предложения, например:

(1) WAK T-1

4 男の人 は 町へ 出掛けて、
Otokonohito wa machi e dekake-te,
мужчина TOP город в выйти-CNV

Мужчина вышел в город...

Впрочем, начальное положение не обязательно, топику могут предшествовать союзные наречия, другие актанты глагола и т. п.:

(2) KEN R-2, 1

ある朝 たらうくん は 目 が 覚めて
Aru asa tarou-kun wa ... (0,6) me ga same-te
один утро Таро-кун TOP глаза NOM открыть-CNV
Однажды утром Таро открыл глаза...

В примере (2) топиальной именной группе предшествует обстоятельство времени «однажды утром», однако функция топики здесь та же, что в примере (1) и принципиально, что пауза после *wa* по типу относится к «паузам внутри клауз».

Однако в корпусе также встречаются примеры, когда именная группа, маркированная показателем *wa*, оказывается не внутри, а на границе клауз. Таких примеров в корпусе всего 18, их можно разделить на 4 типа в зависимости от синтаксической структуры.

В первую очередь положение топиализующей частицы на границе клауз связано с цитационными конструкциями, см. пример (3):

(3) KEN T-1

| | | |
|---------|----------------|-----------------|
| 13 | お子供たち | は |
| | o-kodomo-tachi | wa |
| | HON-ребенок-PL | TOP |
| 14 | 十 円 | 出したら |
| ..(0,3) | juu en | dash-itaru |
| | десять йена | вынимать-COND |
| 15 | 教えて | あげる |
| | oshie-te | age-ru(0,2) |
| | объяснять-CNV | дать-PRS |
| 16 | と | 言いました。 |
| | to | i-imash-ita. |
| | QUOT | сказать-ADR-PST |

Дети сказали: «Если дашь 10 йен, объясним».

Цитируемое предложение фактически представляет собой вложенную предикацию, которая одновременно заполняет валентность глагола «говорить» на прямое дополнение (говорить что?). Цитируемый фрагмент сам по себе представляет сложное предложение из двух клауз. Если «выкинуть» данный фрагмент из главного предложения, то топиализующая частица *wa* окажется в стандартной позиции внутри клаузы.

В пяти из десяти (т. е. в 50%, как и в целом для позиции после *wa*) случаев подобных конструкций с передачей прямой речи возникают паузы после *wa*. Как правило, они довольно краткие, медиана длин данных пауз составляет 0,3 сек., а среднее арифметическое — 0,74 сек. Лишь в одном случае длина паузы достигает 2,7 сек и, видимо, свидетельствует о когнитивных затруднениях говорящего, хотя у этого носителя в целом паузы длиннее, чем у других.

По своей длине паузы после *wa* перед цитацией скорее близки к паузам после *wa* внутри клаузы, чем к паузам на границе клауз (см. подробнее [Комарова, 2008]). Заполненных пауз в данной позиции не встретилось.

На цитирование похож еще один тип конструкций, в которых показатель топики *wa* оказывается в позиции на границе клауз, — это определительные придаточные. В японском языке как в языке с левым ветвлением определительное придаточное предшествует определяемому слову. Оно может разрывать главное предложение и оказываться, как и цитируемый фрагмент, вложенной предикацией, «отрывающей» топик от смыслового глагола.

(4) YAM T-1

| | | |
|----------|------|-----|
| 30 | これ | は |
| ...(0,8) | Kore | wa |
| | Это | TOP |

31 君 の ほしがって いた
 ..(0,2) kimi no hoshigat-te i-ta
 ты GEN хотеть-CNV AUX.PRG-PST
 32 車 だ よ
 kuruma da yo
 машина COP PRT

Это машина, которую ты хотела.

В примере (4) 2-я строка представляет собой относительное придаточное с мишенью релятивизации *kuruma* (машина). Без этого фрагмента главное предложение представляет собой грамматически и семантически законченную предикацию, в отличие от цитационных конструкций, где осталась бы незаполненной валентность глагола говорения на содержание.

Всего в корпусе показатель топики перед определительными придаточными встретился 4 раза. В трех случаях после *wa* были паузы длиной 0,1, 0,2 и 0,2 сек. Заполненных пауз в данной позиции не встретилось.

Длина пауз после топиализующей частицы перед определительными придаточными свидетельствует о том, что данные паузы скорее относятся к «паузам внутри клауз после *wa*», чем к паузам на границах клауз.

Еще один тип конструкций, при котором граница после *wa* совпадает с границей клауз, — это случаи, когда *wa* маркирует союзные имена. При этом сами союзные имена оформляют предикативные обстоятельства, поэтому данный тип границ относится непосредственно к границам между клаузами. Например:

(5) YAM R-2

6 ワイン を 飲んだ 後 は
 ...(0,9) Wain o non-da ato wa
 вино ACC пить-PST после TOP

7 再び スキ に 行く の です が
 futatabi suki ni ik-u no des-u ga
 второй.раз лыжи DAT идти-PRS NML COP.ADR-PRS но

После того, как (он) выпил вина, второй раз пошел на лыжах...

В примере (5) показатель *wa* относится к союзному имени *ato* со значением «после», а в целом *ato wa* присоединяется к финитной форме глагола *non-da* «выпил». Таких случаев в корпусе было 3. В одном из них после топиализующей частицы была незаполненная пауза 0,2 сек.

Данная пауза достаточно коротка, однако описанный тип границ невозможно отнести к типу «внутри клаузы», т. к. союзное имя с показателем топики маркирует глагольную форму, которая в японском языке занимает конечное положение в клаузе.

Наконец, в корпусе встретился один пример, когда показатель *wa* оказался на границе клауз в связи с речевым сбоем:

(6) SHU R-2

1 二つめ の 話 は ==
 futatsume no hanashi wa ==
 второй GEN рассказ TOP
(1,5) ee(0,3) otokonoko ==
 мальчик

Второй рассказ... Мальчик...

В примере (6) говорящий, видимо, хочет сформулировать, о чем будет рассказ в целом, но отказывается от этой идеи и начинает передавать сюжет поэтапно. О когнитивных затруднениях говорящего свидетельствуют довольно длительная незаполненная и заполненная паузы.

Итак, из 18 случаев, когда топиализующая частица оказалась на границе клауз в 14 случаях по типу данную границу можно отнести к «паузам внутри клауз». В 8 случаях из 14, т. е. в 57% случаев на данных границах были паузы. Этот показатель близок к общему показателю «пауз после *wa*» (48%), хотя несколько выше. Заполненных пауз в данных примерах не было, а длина незаполненных, как было показано, также по медиане (0,2 сек) близка к общему показателю для *wa* (0,3 сек).

Оставшиеся 4 примера также не влияют на статистику, т. к. единственная пауза после союзного имени довольно короткая (0,2 сек), а пауза после обрыва — заполненная.

2. Паузы после *ga*

Послелог *ga* — показатель номинатива, маркирующий подлежащее. *Ga* относится к первичным послелогам, т. е. выражает исключительно синтаксическую связь, в данном случае — предиката с первым актантом (см. подробнее [Алпатов, Аркадьев, Подлесская, 2008] стр. 190–194) и может вытесняться показателем топики *wa*.

В корпусе данный послелог встретился 112 раз. В 18 случаях (т. е. в 16%) после него были паузы, лишь одна из них была заполненной, см. пример (7)

(7) MAY R-1, 1

ええと お父さん が ええと 家族 の
 eeto ..(0,2) o-too-san ga ...(0,5) eeto kazoku no
 HEZ HON-отец-сан NOM HEZ семья GEN
 プレゼント を 買い に 行きました。
 purezento o ka-i ni ik-imash-ita
 подарок ACC покупать-CNV DAT идти-ADR-PST

Ну, отец пошел покупать подарок семье...

Так как в корпусе встретился всего один пример с заполненной паузой после *ga*, в данном случае пока-

за предикатом. Все это отражает проблемы планирования данного фрагмента.

В двух случаях послелог *o* оказался в позиции на границе клауза:

(12) YAM R-1, 16

| | | | |
|-----------|---------|-------|-----|
| | 子供 | たち | を |
|(1,3) | Kodomo | tachi | o: |
| | ребенок | PL | ACC |

「母さんは何を||母さんには
 ...(0,8) “kaa san wa nani o|| kaa san ni wa
 мама сан TOP что ACC мама сан DAT TOP

| | | |
|------|-----|-------------|
| 何 | を | 買ったら |
| nani | o | kat-tara... |
| что | ACC | купить-COND |

機嫌を直してもらえる
 ...(0,7) kigen o naosh-ite mora-e-ru,
 настроение ACC починить-CNVполучить-POT-PRS

| | |
|------|------------|
| と | 思う? |
| to | omo-u?” |
| QUOT | думать-PRS |

Детям: «Что бы купить маме, чтобы улучшить (ее) настроение, как вы думаете?»

(13) KEN T-1, 37

| | | | |
|-----------|-------------|---------------|-----------------|
| それで | そういう= | そういう | ことを== |
|(9,2) | Sorede(0,3) | sooyu= | sooyuu koto o== |
| тогда | | вышеназванный | NML ACC |

| | | | |
|------|---|------------------|-----------|
| 家 | へ | 戻って | から |
| uchi | e | modot-te | kara(0,3) |
| дом | в | возвращаться-CNV | после |

Это= После того, как (он) вернулся домой...

В примере (12) *o* вводит прямую речь. Интересно, что глагол говорения (в данном случае, это мог бы быть *tazuneru* «спрашивать») в этом фрагменте опущен. В качестве следующей предикации приводится ответная реплика детей. Как и в примере с *ga* (ср. пример (8)) говорящий делает паузу после *o*, вероятно, чтобы выделить прямую речь. В примере (13) рассказчик отменяет фрагмент перед *o*, с порождением которого у него возникли трудности (в данной клаузе уже был один обрыв), и произносит другой фрагмент, причем он не делает паузы перед откорректированным вариантом.

4. Паузы после *ni*

Послелог *ni* маркирует косвенное дополнение глагола. Он, так же как показатели *ga* и *o* относится к первичным послелогам, однако чаще, чем они, может отражать несинтаксический компонент значения, т. е. выполнять вторичные функции.

Показатель *ni* встретился в корпусе 207 раз. В 48 случаях (т. е. 23%) после него были паузы. 5 из них (т. е. 10%) оказались заполненными. Среднее значение длины пауз после *ni* составило 0,5 сек., медиана — 0,4 сек. и мода — 0,2 сек.

В 8 примерах позиция после *ni* совпала с границей клауза. В первую очередь, как и с другими послелогами, это было связано с началом прямой речи. Интересен пример (14):

(14) YUK T1

| | | | | |
|----------|------------|-----------|---------|----------------------|
| で | とりかか= | あ自分の | 子供 | を= |
| ...(0,7) | De ..(0,4) | torikaka= | ..(0,2) | a jibun no kodomo o= |
| И | | сам | GEN | ребенок ACC |

| | | |
|---------|--|-----|
| ..(0,4) | | に |
| | | ni |
| | | DAT |

| | | | |
|-------|--------------|--------------|------|
| 何 | が | ほしい | か |
| “nani | ga | hoshi-i | ka” |
| Что | NOM | желаемый-PRS | Q |
| を | 一所懸命 | 尋ねる | と |
| o | ishshokenmei | tazune-ru | to |
| ACC | изо.всех.сил | спросить-PRS | TEMP |

Своих детей: «Что вы хотите» — изо всех сил спрашивал...

В данном случае говорящий маркирует адресата локутивного глагола послелогом прямого объекта *o*, однако вспоминает о том, что более грамотно с точки зрения модели управления использовать *ni*. Он отменяет неправильный фрагмент и после паузы заменяет его на более подходящий. При этом паузы после *ni* он не делает, возможно, именно потому, что она возникла перед послелогом (отдельно о паузах перед послелогами будет сказано ниже). Саму же прямую речь как прямой объект глагола говорения говорящий «законно» маркирует послелогом *o*.

Еще один пример аналогично примеру (12) иллюстрирует эллипсис глагола говорения.

(15) SHU T-1

| | | | | |
|---|-----------|-------|--------------|-----|
| 6 |(2,4) | で | 子供たち | に |
| | | de | kodomo-tachi | ni |
| | | Тогда | ребенок-PL | DAT |
| 7 |(1,0) | | 子供たち | は |
| | | | kodomo-tachi | wa |
| | | | ребенок-PL | TOP |

Тогда к детям... Дети...

В примере (15) описывается ситуация, когда отец обращается к детям и дальше следует их прямая речь. Однако говорящий опускает локутивный предикат, не заканчивая 6-ю строку. В данном случае непонятно, насколько серьезные когнитивные затруднения возникли у рассказчика. То ли он не смог закончить фразу и подобрать нужный глагол, то ли ре-

шил, что и так ясно, что он имел в виду, и заканчивать данное предложение не обязательно. Незаполненная пауза перед следующей репликой и отсутствие попыток произнести какой-нибудь глагол говорения свидетельствуют в пользу второго предположения.

Наконец, интересный пример, когда показатель *ni* оказался не внутри клаузы, связан со случаем парцелляции:

(16) KIM T-2

- | | | | |
|---|-------------------|-----------|----------------|
| 3 | 紅茶 と | ピザ を | 食べました。 |
| | koocha to ..(0,3) | pizza o | tabe-mash-ita. |
| | черный.чай и | пицца ACC | есть-ADR-PST |
| 4 | 朝ご飯 | に。 | |
| | Asagohan | ni. | |
| | Завтрак | DAT | |

Съел пиццу с черным чаем. На завтрак.

В примере (16) говорящий выносит обстоятельство за глагол, хотя в японском языке глагол должен занимать конечное положение в предложении. Это может быть связано с тем, что рассказчик забыл нужный элемент и не успел поставить его перед глаголом. Или он хочет коммуникативно выделить этот фрагмент, например, указывая, что пицца не самый стандартный вариант завтрака.

Всего в корпусе встретилось 7 случаев, когда показатель *ni* использовался для введения прямой речи. При этом, кроме случая, рассмотренного в примере (15), где пауза после *ni* составила 1,0 сек., еще трижды перед прямой речью возникали короткие паузы (0,1, 0,2 и 0,2 сек.).

В примере (16) пауза после парцелированного элемента составила 1,0 сек. По своим длине и позиции она относится к типу «пауз на границе предложений».

5. Паузы после *no*

Послелог *no* маркирует, в отличие от рассмотренных ранее *ga*, *o*, *ni*, атрибутивную, а не актантную связь. Он встретился в корпусе 96 раз. Всего в 4-х случаях (т.е. в 4%) после *no* были паузы, причем две из них были заполненными.

(17) ISI R-2, 26

- | | | | |
|-----------------|---------------|-----------------|--|
| 私 の ええと | 左 腕 と | 右 足 は | |
| watashi no eeto | hidari ude to | migi ashi wa | |
| я GEN HEZ | левый рука и | правый нога TOP | |
- Мои левая рука и правая нога...*

В примере (17) проблемы говорящего маркирует лексический показатель хезитации *eeto* (аналог русского «ну»). Видимо, когнитивные затруднения

рассказчика связаны с использованием таких понятий, как «правый» и «левый», и он задумывается, пытаясь по картинке понять, о каких именно руке и ноге идет речь.

Приведем еще один пример с паузой после *no*:

(18) YAM T-1, 23

- | | | | |
|---------------------|--------------------|---------|-------------|
| こ ち ら の | 車 の | 方 け= | え |
| ...(2,5) Kochira no | ...(0,9) kuruma no | hoo | ...(0,5) k= |
| Здесь GEN | машина GEN | сторона | |

Это... Эта машина...

В примере (18) пауза после *no* — незаполненная, однако ее длина и наличие коррекции в последующем фрагменте свидетельствуют о когнитивных затруднениях говорящего. Еще в одном случае незаполненная пауза после *no* составила 0,1 сек. За одну именную группу перед ней была заполненная пауза, что также позволяет предположить некоторые трудности с порождением текста у рассказчика, однако, длина паузы незначительна, так что доказать это невозможно.

Таким образом, среднее арифметическое для длин данных пауз составило 0,8 сек., медиана — 0,9 сек., однако эти данные не показательны, т. к. высчитывались на основе 3-х примеров (четвертая пауза была заполненной, без незаполненного или долежисически заполненного фрагмента, а для лексических заполнителей время не высчитывается).

6. Паузы перед послелогом

Отдельно хочется остановиться на позиции перед послелогом. В корпусе всего один (!) раз встретилась пауза перед одним из рассмотренных послелогов, см. пример (19)

(19) ISI T-2, 23

- | | | | |
|------------|-----------|----------|---------|
| 左 腕 | を | 骨 折 | して |
| hidari ude | ..(0,2) o | kossetsu | shi-te |
| левый рука | ACC | перелом | VRB-CNV |
- (Он) сломал левую руку...*

Еще в одном случае говорящий тянул слово перед послелогом *ni*. В примере (19) нет речевых сбоев и пауза перед послелогом достаточно краткая, однако ее исключительно малая вероятность (это единственная пауза на 913 вхождений послелогов) позволяет говорить о возможных проблемах планирования.

Итак, данные корпуса свидетельствуют о том, что паузы перед послелогом практически невозможны.

Что касается позиции после послелогов, то:

- 1) Паузы после *no* возникают всего в 4% и в 50% случаев — заполненные, т. е. почти всегда свидетельствуют о хезитации;

- 2) Паузы после *o* возникают в 7,5% случаев, в 18% случаев заполненные, таким образом, вероятность этих пауз также минимальна;
- 3) Паузы после *ga* возникают в 16% случаев, в лишь 5% случаев они заполненные, что свидетельствует о возможности данных пауз;
- 4) Паузы после *ni* возникают еще чаще — в 23% случаев, в 10% случаев они заполненные, т.е., видимо, паузы в данной позиции «законны».
- 5) Наконец, паузы после *wa* встречаются в 48% случаев. В 27% случаев они бывают заполненными.

Итак, частота пауз в позиции после топики и вероятность наличия заполнителя указывают

на то, что эта позиция является одной из «позиций планирования», как, например, граница предложений (где паузы встречаются в 94% случаев, и 23% из них — заполненные, см. [Комарова, 2008]). Это лишний раз подтверждает тот факт, что топик относится ко всему предложению в целом, т. е., говорящий может сформулировать, о чем будет идти речь, но еще не знает, что именно и как он скажет.

Кроме того, данные по послелогам *ga*, *ni*, *o* могут свидетельствовать о существовании некоторой иерархии свободы актантов. Так, наименее вероятны паузы после *o*, несколько выше их вероятность после *ga* и, наконец, после *ni* они уже довольно частотны.

Литература

1. Алпатов В. М., Аркадьев П. М., Подлеская В. И. Теоретическая грамматика японского языка, М. 2008
2. Кибрик А. А., Подлеская В. И. (ред.) Рассказы о сноведениях: корпусное исследование устного русского дискурса. Москва, 2009.
3. Комарова А. Д. Паузация в японском языке на границах синтаксических единиц разного уровня: корпусное исследование // международная конференция Диалог 2008
4. Кривнова О. Ф., Чардин И. С. Паузирование при автоматическом синтезе речи // Теория и практика речевых исследований (АПСО-99). Москва, 1999.
5. Chafe. W. Discourse, consciousness, and time. Chicago, 1994.
6. Iwasaki S. The Structure of the Intonation Unit in Japanese // Japanese/Korean Linguistics, vol. 3, ed. by Soonja Choi. Stanford, 1993, 39-53
7. Sakura C., Fuji S. Intonation units, information structure, and grammatical constructions in Japanese and English. // Oral presentations. University of Tokyo, 2006.