

ВИЗНАЧЕННЯ МЕЖ СКЛАДНИХ РЕЧЕНЬ У СУЧАСНІЙ КИТАЙСЬКІЙ МОВІ

Під час проведення мовознавчих досліджень, що стосуються усного мовлення, часто виникає потреба транскрибування аудіо-інформації засобами письма тієї чи іншої мови. Залежно від мовних рівнів, на які спрямовані ці дослідження, можуть виникати різноманітні перешкоди та проблеми при описі та класифікації деяких граматичних конструкцій. Зокрема, при спробі синтаксичного розбору фрази усного, найчастіше спонтанного, мовлення перед аудитором неодмінно постає питання розмежування фрази на окремі речення, принаймні для суто формального позначення кінця та початку різних речень. Звісно, це питання особливо гостро стоїть у контексті китайської мови, де навіть формальні ознаки межі слова відсутні та, в певному розумінні, не передбачені діючою наразі системою письма.

Потреба у відносно чітко визначених способах синтаксичного поділу мовленнєвого потоку виникає на фоні швидкого розвитку найбільш актуальної у китайській лінгвістиці галузі – корпусної лінгвістики, що останнім часом здебільшого спрямована на створення масштабних корпусів, серед яких найбільші засновуються саме на аудіо- та відео-записах (Spoken Beijing Mandarin BJKY, TV and Radio transcripts in Broadcast Media Language Corpus, Child and Family Interaction Discourse, Wenzhou Spoken Corpus, Sinica Balanced Corpus, NCCU Corpus of Spoken Chinese і т. д.).

Важливість залучення матеріалів спонтанного усного мовлення складно переоцінити, коли йдеться про вивчення фундаментальних характеристик мови, оскільки саме в ньому знаходять своє відображення такі тонкі нюанси, котрі практично неможливо відслідкувати у рафінованій писемній мові. Наразі саме у створенні корпусів, що базуються на транскрипціях аудіо- та відео-записів, вбачають найбільші перспективи китайські мовознавці. Тому, на відміну від більшості сучасних корпусів китайської мови, або китайських мов (до них можна умовно віднести синографічні мови – кантонська, тайванська, сингапурська тощо), де частка усного мовлення зовсім незначна та, крім того, взята з

писемних джерел, такі масиви транскрибованого і відповідно впорядкованого усного мовлення можуть дати цілком нові можливості для практичного застосування.

Разом із тим у процесі створення корпусів, значною частиною або повністю заснованих на усному мовленні, виникають деякі суперечливі моменти. Починаючи від етапу відбору інформації, та закінчуючи безпосереднім формуванням інтерфейсу роботи з корпусом. Наприклад, нерідко перед дослідниками постає питання вибору засобів токенизації [Tseng S.-C. 2004, 93], або необхідності створення нового програмного забезпечення, що було б більш пристосованим для роботи з матеріалом конкретної мови [Chui, Lai 2008, 120].

Неодноразово дослідниками синтаксису сучасної китайської мови порушувалося питання про визначення меж речень та їхню класифікацію за типами синтаксичного зв'язку [Gao 1986, Горелов 1963, Lü 2002]. Неоднозначною є проблема принципів членування складного речення на менші підрядні або сурядні складові. Здебільшого питання зводиться до того, що ж вважати найбільш важливою конструктивною ознакою складного речення: єдину загальну синтаксичну структуру чи завершеність інтонаційного вираження окремого уривка.

Формальний підхід до вирішення цієї дилеми, тобто вважати окремим реченням, складним або простим, те, що у вигляді тексту знаходиться між двома крапками, підходить винятково для поверхових досліджень і зовсім не відповідає вимогам дослідження усного мовлення, оскільки межі речення в усному матеріалі можуть сильно відрізнитися від фактичної пунктуації при транскрибуванні. Такий підхід особливо неприйнятний у випадках дослідження спонтанного мовлення, специфіка якого полягає, зокрема, у неоднорідності синтаксичного впорядкування висловів та в орієнтації мовця на досягнення інтонаційної завершеності [Багмут 1977, 35].

Навіть нехтуючи специфікою спонтанного мовлення та орієнтуючись на розгляд підготованого мовлення, неважко помітити, що для реципієнта у процесі комунікації визначну роль відіграє саме інтонаційний вигляд вислову та його відповідність структурі та змістові повідомлення [Бубнова 1991, 203]. Щоправда, для мовця в деяких випадках можливі різні ступені спонтанності у

реалізації інтонаційної чи синтаксичної завершеності вислову з огляду на умови, за яких проходить спілкування. Цілком закономірно, що при психологічно зрівноваженому та спокійному оточенні можливе до певної міри відхилення від загальноприйнятих правил оформлення мовлення, а відповідно проявляються деякі характерні явища порушення мовленнєвого потоку (хезитація, випущення слів тощо) як на рівні граматичному, так і на рівні фонологічному чи інтонаційному.

Виходячи з вищеназваних причин, цілком доречним є таке зауваження: “під складним реченням слід мати на увазі складну синтаксичну одиницю, що складається із двох чи більше частин, котрі утворюють єдине змістове та інтонаційне ціле” [Горелов 1963, 9].

Дещо інакшого положення дотримувався Гао Мінкай. На його думку, в китайській мові відсутня морфологія, тому морфологічної характеристики службових слів він не дає, а лише визначає їхні функції відповідно до тих зв’язків, котрі вони позначають у реченні. Слова, що у нинішній класифікації сучасної китайської мови прийнято називати службовими словами з функціями сполучників (тобто 也, 就, 才, 便 тощо), Гао Мінкай називає просто зв’язковими словами. Разом із тим, беручи до уваги відмінність цих слів від низки інших, котрі він також відносить до зв’язкових, він називає їх спеціальним терміном – “承词”. Цей термін слід розуміти таким чином, що слова цього типу встановлюють зв’язок між попередньою та наступною частинами речення [Gaо 1986, 256]. Натомість Ляо Шуцянь відносить речення зі словами вказаного типу до безсполучникових, при цьому фактично зараховуючи їх до засобу зв’язку між частинами речення.

Такий підхід уникає залучення інтонаційної інформації речення при його поділі на частини, тобто засновується винятково на текстовій інтерпретації мовлення, що не відповідає вимогам, що можуть постати при розгляді усного мовлення. З іншого боку, він також нехтує семантичною складовою вислову, просто випускаючи її з уваги, що зводить його підхід до доволі одностороннього аналізу матеріалу.

У межах сучасної китайської мови можна відмітити два основні способи поєднання частин складного речення: асиндетичний та синдетичний. Асиндетичний спосіб – це такий синтаксичний

прийом, за якого частини складного речення поєднуються за допомогою двох обов'язкових для кожного складного речення засобів: інтонації та порядку положення частин [Горелов 1963, 20].

При синдетичному способі, окрім інтонації та порядку частин речення, використовуються також один або кілька спеціальних формальних засобів синтаксичного зв'язку. При асиндетичному способі змістові відношення між частинами складного речення в ряді випадків проявляються недостатньо повно. Так, наприклад, іноді буває складно розділити відношення часу та умови. У цьому розумінні синдетичний спосіб є значно більш довершеним засобом синтаксичного зв'язку.

Серед загальних особливостей вжитку сполучників у складних реченнях слід також вказати одночасне використання сурядних сполучників та часток сполучникового типу. Ця особливість обумовлена сукупністю тих умов, що виникають у складних реченнях взагалі та у складносурядних реченнях зокрема, під впливом взаємодії різноманітних засобів синтаксичного зв'язку [Gao 1986, 139].

Там само автор неодноразово вказує на парне та одиночне функціонування сполучників та сполучникових часток. Разом із тим у ряді випадків неможливо чітко встановити, які саме сполучники були випущені та скільки їх було або може бути в конкретному реченні. Відповідно, важливим із точки зору розгляду міжрівневих взаємодій при асиндетичному зв'язку є зіставлення не тільки інтонаційних контурів та синтаксичних конструкцій із відсутніми формальними засобами сполучення як такими, але й окремі випадки часткового вжитку таких сполучникових пар, де наявна лише одна з частин.

Доволі проблематично чітко встановити, в якому вигляді функціонують деякі сполучники – попарно чи поодинокі. Проте значна частина таких граматичних конструкцій, що іноді також називаються “рамковими”, доволі широко описані у граматиках сучасної китайської мови та мають цілком визначені та закономірні принципи функціонування. З іншого боку, можна лише теоретично припустити подальшу їхню долю в історичних перетвореннях мови, оскільки вже на сьогодні можна стверджувати про деякі скорочення серед них. Так, наприклад, практично зник з ужитку парний сполучник “不错...但是”, де редукувалася перша частина.

Скорочення таких складних сполучникових конструкцій у сучасному китайському мовленні, можливо, свідчить про тенденцію до їхнього поступового редукування до рівня простих односкладових сполучників. Наявні свідчення з веньяню, де переважна більшість службових слів взагалі була моносилабічною, насправді суперечать таким припущенням, але слід врахувати також, що в означеному випадку йдеться про усне мовлення, котре мало докорінні розбіжності із класичною писемною мовою вже у 17-му столітті. Тому цілком імовірно, під впливом еволюційних процесів в усному мовленні дещо змінилися закономірності вжитку сполучників та сполучникових часток, що і без того не були унормовані.

Отже, основними факторами визначення певного вислову як єдиного цілого є завершеність із точки зору змісту, інтонації та синтаксису. Асиндетичний тип зв'язку між частинами складного речення, що відповідає вищеназваним умовам завершеності в плані семантики, синтаксису та інтонації, при своїй формальній невираженості може спостерігатися на рівні із синтетичним зв'язком. Оформлення ж асиндетичного зв'язку відбувається саме за допомогою засобів інтонації. Крім того, складні асиндетичні речення, окрім синтаксичного вигляду, відрізняються між собою за типом тих сполучників чи сполучникових часток, що були опущені у процесі продукування. Наприклад, так звані “парні” сполучники чи “рамкові конструкції”, що виконують роль сполучників, можуть бути наявними у асиндетичних реченнях тільки частково, що відповідним чином видозмінює і структуру інтонаційного контуру речення.

З іншого боку, при можливості проведення аналізу як усного, так і писемного мовлення слід мати на увазі, що ці два різновиди одного за суттю процесу насправді докорінно відрізняються як у засобах творення, так і у призначеннях. Отже і критерії до опису такого матеріалу мають відповідним чином відрізнятися.

Разом із тим, у більшості наведених вище прикладів проводиться аналіз саме мови писемної, хоч і йдеться про “інтонаційну завершеність”, тоді як в усному мовленні немає настільки ж очевидних формальних засобів визначення меж речення, як розділові знаки, котрі мають відображати зміни інтонації на письмі. Чи завжди пунктуація є послідовною та однозначною при

встановленні меж речення, наприклад, під час транскрибування усного мовлення?

Якщо розглянути деякі приклади, наведені у статті Сю Яньцін [Hsu 2010, 14–16], виникають питання щодо послідовності розмежування речень у репліках із діалогів під номерами 18 та 20. І хоч можливість прослухати оригінальний аудіо-матеріал відсутня, але цілком очевидно, що в першому випадку аудитор керувався саме інтонаційним виглядом фрази:

“我是赞成是每天给他零用钱这样，但是，不要太多，多了也许会造成他们额外的动作出来[...]”

При розборі цього прикладу на синтаксичні складові можна з великою імовірністю припустити, що темп мовлення був занадто швидким для того, щоб при транскрибуванні замість одного складного речення були записані кілька. При тому, що для більшої зручності у сприйнятті такого тексту читачем важливо не тільки доцільне використання розділових знаків, а й видалення повторів та самокоментувань мовця.

У іншому прикладі з цієї статті можна помітити цілком граматично вірне використання “переверненої” коми в такому випадку, де, найімовірніше, в ній формальної потреби (а саме розділення однорідних членів речення) немає:

对、对，社会责任。

Таким чином, ми бачимо, що навіть у випадках, коли проводиться синтаксичний чи лексичний аналіз записів усного мовлення, при транскрибуванні можуть допускатися неоднозначні та неоднорідні використання пунктуаційних знаків. Що характерно, загальноприйняті на китайському телебаченні китайськомовні субтитри до практично будь-яких телепередач (окрім прямих трансляцій) взагалі записуються без ком та крапок, лише із пробілами в тих місцях, де у мовленні звучить доволі тривала пауза, або де є потреба розділити очевидно різні за змістом речення. Крім того, цілком допустимо та рекомендовано спрощувати фактичні репліки мовців, уникаючи повторень та граматичних помилок мовців.

З іншого боку, на прикладі анотування спонтанного мовлення англійської мови та опису повного процесу його проведення [Heeman, Allen 1999, 529] очевидно, що у кінцевому результаті більше уваги приділено розмежуванню речень за синтаксичною

цілісністю, тоді як відсутністю пауз у деяких випадках знехтувано. З іншого боку, анотування та транскрибування, що описується у цій статті, спрямоване на досягнення дещо інших цілей – дослідження можливостей моделювання спонтанного мовлення у доволі вузьких тематичних рамках.

Таким чином, незалежно від потреб, якими керується аудитор, котрий проводить анотацію, можна умовно виділити два підходи у проведенні транскрибування усного мовлення з точки зору використання пунктуаційних засобів для визначення меж речень: суто інтонаційний, за якого допустимим до певної міри є нехтування синтаксичним поділом мовленнєвого потоку, та більше спрямований на синтаксично послідовне та граматично обумовлене членування мовленнєвого потоку. Варто зауважити, що для деяких корпусів усного китайського мовлення, наприклад, для створеного в Китайському Університеті інформаційно-комунікаційних технологій (TV and Radio transcripts in Broadcast Media Language Corpus), характерні тенденції до більш узгодженого із чинними правилами правопису оформлення тексту пунктуаційними знаками, на відміну від тих корпусів, де розділові знаки вживаються рідше та майже виключно для позначення пауз у мовленні. Проте така специфіка може бути обумовлена тим, що значну частину корпусу складає усне мовлення, що не є спонтанним, а в більшості випадків підготованим або й начитаним. Разом із тим, аналогічно насиченою розділовими знаками є й та частина транскрибованих записів, що належать до спонтанного мовлення (діалоги та інтерв'ю).

При інтонаційно спрямованому розмежуванні речень застосовують різні засоби формального членування – як звичайні розділові знаки, наприклад, коми, крапки, тире тощо для позначення пауз, так і специфічні позначки. Введення особливих відміток для різного роду пауз, що є типовими саме для спонтанного усного мовлення, поєднується також і з використанням звичайної пунктуації. Втім, розрізнення між хезитаційними та іншими паузами немає.

Виходячи з матеріалів тих корпусів усного мовлення, що перебувають у вільному доступі, можна зробити висновок, що питання визначення меж речень не завжди є першочерговим та часто лишається поза увагою авторів. І через це немає однозначного

рішення щодо використання формальних позначень розмежування цілісних та незалежних одне від одного речень. Пунктуація може бути показником як паузального, а отже й інтонаційного членування речення, так і синтаксичного поділу на складові частини. Незважаючи на таку невизначеність, подібні корпуси є потужним засобом аналізу таких проявів мовлення, що реалізуються саме у процесах спонтанного говоріння. У тому ж разі, коли корпус містить транскрипції підготованого або начитаного усного мовлення, суттєвих відмінностей від текстових корпусів практично немає.

У цілому при транскрибуванні спонтанного мовлення більш доречно орієнтуватися саме на інтонаційний вигляд та паузальне членування речення, оскільки таким чином більше уваги приділяється нелінійності та неоднорідності, котрі є важливими характеристиками усного непідготованого мовлення. Відповідно до потреб, що висуваються у процесі обробки матеріалу, та до завдань, що були поставлені перед експериментатором, слід використовувати той спосіб запису та токенизації, що буде найоптимальнішим у конкретних умовах. Наприклад, при необхідності аналізу саме спонтанного мовлення варто брати за основу інтонаційний принцип членування, тоді як для підготованого або начитаного мовлення розмежування фрази на синтаксичні складові може збігатися з межами більших просодичних одиниць.

ЛІТЕРАТУРА

- Багмут А. Й. **Типологія інтонації мовлення**. Київ, 1977.
- Бубнова Г. И., Гарбовский Н. К. **Письменная и устная коммуникация. Синтаксис и просодия**. Москва, 1991.
- Горелов В. И. **Союзы в сложном предложении современного китайского литературного языка**. Москва, 1963.
- Chui Kawai, Lai Huei-Ling. The NCCU Corpus of Spoken Chinese: Mandarin, Hakka and Southern Min // **Taiwan Journal of Linguistics**. Vol. 6.2, Taipei: Crane Publishing, 2008.
- Gao Ming-kai. **Hanyu yufa lun**. Beijing: Shangwu yinshuguan, 1986.
- Heeman P. A., Allen J. F. Speech repairs, intonational phrases, and discourse markers: modeling speakers utterances in spoken dialogue //

Journal of Computational Linguistics. Vol. 25 Issue 4, MIT Press Cambridge, December 1999.

Hirschberg Julia, Litman Diane. Empirical studies on the disambiguation of cue phrases // **Journal for Computational Linguistics.** Vol. 19, Issue 3, MIT Press Cambridge, September 1993.

Hsu Yen-Ching. General noun, ‘dongxi’, in spoken Mandarin // **NCL, 2010.**

Lü Shu-Xiang. Hanyu yufa fenxi wenti // **Lü Shu-Xiang quanji, di er juan.** Shenyang: Liaoning jiaoyu chubanshe, 2002.

Tseng Chiu-yu. Prosody Analysis // **Advances in Chinese Spoken Language Processing.** Singapore: World Scientific Publishing, 2007.

Tseng Shu-Chuan. Processing Mandarin Spoken Corpora // **TAL.** Vol. 45, 2004.

Tseng Shu-Chuan. Spoken corpora and analysis of natural speech // **Taiwan Journal of Linguistics.** Vol. 6.2, Taipei: Crane Publishing, 2008.

Tseng Shu-Chuan. Toward a large spontaneous Mandarin dialogue corpus // **SIGDIAL’01 Proceedings of the Second SIGdial Workshop on Discourse and Dialogue.** Stroudsburg, Vol. 16, Association for Computational Linguistics, 2001.