

## ТРАНСКРИПЦІЯ Й ТРАНСЛІТЕРАЦІЯ АРАБОМОВНИХ ТЕКСТІВ: ВІД ДІАКРИТИКИ ДО ПРОСОДИЧНИХ ШАБЛОНІВ

**З**вукові явища конкретної мови, як сегментного, так і супрасегментного рівня, представлені у вигляді окремих фонем і просодичних одиниць, визначають ту систему фонетичного запису звукового ладу мов світу (транскрипції), яка вважається стандартною в сучасній лінгвістиці.

Фонетичною транскрипцією (від лат. *trans* ‘через’ і *scribere* ‘писати’) зазвичай називають особливу систему письма, що використовується для точного передання звукового складу усного або письмового мовлення. Залежно від способу запису й від того, які звукові одиниці є предметом транскрипції, розрізнюють власне фонетичну, фонологічну й практичну транскрипцію.

Фонетична транскрипція має на меті досягнення повної відповідності зі звучанням і враховує усі аллофони, комбінаторні й позиційні варіанти фонем, одиниці супрасегментного рівня, реалізовані в мовленні, відтак вона є найбільш точною та універсальною.

Фонематична транскрипція використовується для передання фонемного складу слова або морфеми, відображає тільки ті особливості мовлення, які мають смислорозрізнювальне значення, й залежить від визначення фонем і аллофонів у межах конкретної системи транскрипції.

Практична транскрипція використовується для передання одиниць мовлення однієї мови стандартними засобами іншої, цілком визначається графічною системою останньої і тому не повинна містити, на відміну від двох попередніх, додаткових знаків.

Від транскрипції треба відрізняти транслітерацію (від лат. *trans* ‘через’ і *littera* ‘літера’), яка являє собою політерне передання мовних одиниць усіх рівнів (до тексту включно), за-писаніми засобами однієї графічної системи, засобами іншої. Базуючись на конкретному алфавіті, припускає умовне використання літер, а також введення додаткових діакритичних знаків.

Найвідомішою й найбільш розробленою та стандартизованою системою транскрипції є система МФА (Міжнародний фонетичний алфавіт), прийнята у 1889 році. Важливою перевагою класифікації МФА є те, що вона за тривалий час свого використання в ролі універсальної набула рис міжнародного стандарту. Тому, за умов коректного використання, вона дає можливість описувати фонетичні явища як з метою практичного вивчення конкретної мови, так і з метою ознайомлення дослідників, у тому числі й із суміжних галузей, з цими явищами. Зрозуміло, що через свою універсальність використання цієї системи для конкретної мови вимагає певної адаптації. Варіант, прийнятний для АЛМ, можна знайти, наприклад, у [Handbook... 2005, 51].

Головним недоліком цієї системи є обмеження артикуляційного та акустичного розмаїття звукової системи конкретної мови фіксованим набором характеристик, без чого формування ознак-класифікаторів, на розрізенні яких базується будь-яка система класифікації, є неможливим.

Основним набором символів, що їх використовує система МФА, є латинська графіка. Зрозуміло, що фонем у мовах світу значно більше, ніж латинських літер, тому задля досягнення універсальності необхідно застосувати діакритичні позначки та спеціальні символи, модифіковані форми літер (великі, перевернуті літери, “хвостики” й “гачеки”, неалфавітні символи тощо). Символи для позначення одиниць супрасегментного рівня не надто стандартизовані й тому являють собою найбільш дискусійну частину МФА.

Інша система транскрипції склалася в дослідженнях з порівняльно-історичного мовознавства. Основними відмінностями тут є позначення африкат одним символом замість двох,

ефективних – за допомогою крапки під літерою, а сибілянтів – за допомогою додавання “гачка”. Відмінності стосуються також позначення деяких голосних [Кодзасов... 2001, 319]

Набір диференційних ознак складає основу також і для двох вищезгаданих типів запису – фонетичного й фонематичного (фонологічного), розрізнення яких було введено в обіг Л.В. Щербою [Зиндер 1979, 291].

У фонетичному записі звуки позначаються за допомогою спеціальних символів, хоча за основу й беруться літери якогось конкретного алфавіту (у випадку МФА це латина), а у фонематичному зазвичай використовується символ, який позначає основний варіант фонеми, що реалізується у сильній позиції або взагалі за відсутності контексту.

Обидва типи запису використовуються як у практичних, так і в дослідницьких та навчальних цілях, і обидва вони мають базуватися скоріше на міжмовних, ніж на внутрішньомовних міркуваннях.

Справді, внутрішньомовна транскрипція має бути використаною лише для однієї мови і тому обмежена фонематично, без ідеї порівняння з іншими мовами. Але навіть у цьому випадку знаків алфавіту цієї мови для точного передання її звукової системи з усіма припустимими у цій мові алофонами явно не вистачить, адже жодна система письма, крім спеціальних фонетичних алфавітів, не передбачає точне фонетичне або фонематичне відтворення звукової системи. Отже, використання додаткових знаків у системах транскрипції є неминучим.

Водночас кількість теоретично можливих звуків, використовуваних у мовленні, можна вважати безкінечною, тому будь-яка система фонетичної транскрипції неминуче припускається спрошені. Отже, універсальні системи, такі як МФА, передаватимуть скоріше не точне звучання, а співвідношення між фонемами та алофонами тієї чи іншої мови. Разом із тим фонетична транскрипція, яка за визначенням є значно близчкою до реального звучання, ніж фонематична, є можливою тільки тоді, коли окрема фонема інтерпретується як один із завданої кількості елементів фонематичного простору. Отже, при переданні звукової системи мови на письмі “спрошення реальності” є неминучим. З цієї точки зору усі системи письма, створені людством, можна охарактеризувати за релевантністю щодо звукового боку мовлення, яке вони фіксують. Найменш релевантним у такому представленні буде, очевидно, ідеографічне письмо, найрелевантнішими – фонетичні алфавіти. Решта систем графічного передання мовлення займатиме проміжне положення.

Ще одним завданням, яке виникає при письмовій фіксації звуків мовлення засобами алфавітного письма, що створювалося для іншої мови, – а така ситуація автоматично виникає в усіх неарабомовних публікаціях, присвячених арабській мові, – є транслітерація, тобто передання тексту, записаного арабською графікою, засобами іншої графіки. Найчастіше це латиниця, а у вітчизняних публікаціях часто можна зустріти кириличну транслітерацію. Основний принцип “ідеальної” транслітерації полягає у застосуванні одного символу графіки, за допомогою якої здійснюється транслітерація, для передання одного символу оригінальної графіки. Зрозуміло, що за дотримання цієї вимоги у випадку як латиничної, так і кириличної транслітерації арабомовного тексту неможливо обйтися без додаткових знаків. Тут теж можливими є два основних підходи: фонетичний (скоріше, фонематичний) і формальний.

Перший підхід, який можна назвати “фонетичною транслітерацією”, багато в чому нагадує фонематичну транскрипцію, але є набагато спрощенішим і схематичнішим. Утім, за його допомогою можна задовільно передавати “ідеальні” прочитання арабських літер без урахування алофонів. Майже ігноруються за такого підходу й одиниці супрасегментного рівня, адже у випадку КАМ і АЛМ вони визначаються правилами самої мови й система транскрипції, що є адекватною до звукової системи мови, повинна дозволяти позначати ці одиниці за допомогою стандартних засобів сегментного рівня. Арабські фонеми, що сильно відрізняються від “стандартних” латинських, позначаються за допомогою відповідної діакритики. Стандартів такої транслітерації не існує, але в навчальній і спеціальній літературі з арабістики склалася певна традиція, якої, з незначними відмінностями, дотримується більшість арабістів (див. табл. 1).

На жаль, не всі символи з діакритичними позначками, наведені у табл. 1, входять до стандартного набору Unicode. Тому у деяких випадках через технічні причини замість відсутніх символів використовуються спеціальні літери, що мають накреслення, подібне до необхідного, і вживаються в алфавітах мов світу на позначення фонем, які можуть не

мати нічого спільного з відповідними арабськими фонемами [див., напр.: Рыбалкин 2000, 4–5]. Головним принципом такої практики є використання символів, що входять до різних наборів Unicode, для позначення арабських фонем, відсутніх у мовах, яким відповідають ці набори, наприклад  $\hat{h}$  для  $\zeta$  тощо. Таку транслітерацію можна назвати формальною, так само як і ту, що використовується там, де навіть стандартна діакритика є небажаною або технічно неможливою. Прикладом такої транслітерації є передання арабського тексту в SMS-повідомленнях, чаті, а також системи Arabizi і Arabesh (див. табл. 1). Як можна побачити, тут крім стандартних латинських літер використовуються цифри. Іноді така транслітерація зустрічається в англомовних публікаціях з арабістики й суміжних дисциплін. Прикладом формального підходу в кириличній транслітерації є використання літери “г” на позначення літери  $\zeta$  в україномовних виданнях. Прагненням уникнути спеціальних символів у транслітерації пояснюється також використання диграфів, що є відходом від основного принципу транслітерації й може привести до неоднозначного прочитання (наприклад,  $sh$  для  $\zeta$ ).

Ще більш спрощеною, хоча й значно менш формальною, є транслітерація, яка використовується для передання власних імен, географічних назв, специфічних арабських реалій тощо в нелінгвістичних текстах. Тут оригінальне звучання передається вельми приблизно за допомогою лише стандартних засобів мови основного тексту. Такий підхід є близьким до практичної транскрипції. При цьому перевага часто віддається діалектним варіантам замість літературних – наприклад, “Шарджа” замість “Шаріка” для  $\text{شَارِقَة}$ .

Існують також міжнародні стандарти й нормативні акти транслітерації географічних назв і власних імен (напр., UNGEGN, BGN/PCGN, ACGN тощо), але такі системи постійно вдосконалюються, й тому створення єдиної універсальної системи транслітерації на їхній основі навряд чи є можливим.

Окрім того, варто розрізнювати мовний контекст, у якому функціонує елемент, для якого здійснюється транскрипція, а особливо транслітерація. Так, якщо йдеться про АЛМ або про її діалекти, більшість систем транскрипції транслітерації орієнтуватимуться на реалізацію відповідних фонем саме у цій мові. Але якщо транслітерується арабське слово, запозичене іншою мовою, наприклад турецькою або перською, можливе суттєве відхилення від арабського звучання. Наприклад, через відсутність відповідних фарингальних та фарингалізованих приголосних у цих мовах такі власні імена, як  $\text{محمد}$  і  $\text{حسین}$ , будуть передаватися як Muhammad і Husayn замість Muhammad і Ḥusayn відповідно. Цей принцип може поширюватися й на практичну транскрипцію – тут варто згадати практику написання цих імен як “Мугаммад” і “Гусейн” в україномовних текстах. На нашу думку, цей підхід є хибним принаймні з фонологічної точки зору, адже опозиція, глухість / дзвінкість, є однією з найголовніших корелятивних опозицій фонологічної системи української мови. Відтак, коли йдеться про транскрибування цих імен, якщо їх вжито в арабомовному контексті, арабський глухий фарингальний звук “ $\hat{h}$ ”, позначуваний літерою  $\zeta$ , варто передавати за допомогою глухого ж велярного українського звука “ха”, а не дзвінкого нижньо-фарингального “га”. Разом із тим останній варіант є доцільним, коли йдеться про перський або турецький мовний контекст.

У таблиці 1 наведено найбільш уживані символи транскрипції й транслітерації арабських текстів і фрагментів на основі латиниці й кирилиці.

Нижче наведено список основних систем латинської й кириличної транскрипції й транслітерації арабомовних текстів і фрагментів із зазначенням там, де це можливо, часу їхнього виникнення:

МФА (IPA) (1889) – Міжнародний фонетичний алфавіт (International Phonetic Alphabet). Система транскрипції, прийнята Міжнародною фонетичною асоціацією (IPA – International Phonetic Association).

DMG – Німецька східна асоціація (Deutsche Morgenländische Gesellschaft). Систему транслітерації, прийняту у межах Міжнародної конвенції сходознавців у Римі в 1936 р., розроблено К. Брокельманом і Г. Вером (Carl Brockelmann, Hans Wehr). Цю систему взято за основу в авторитетному словнику Ганса Вера (Arabisches Wörterbuch für die Schriftsprache der Gegenwart (1952), перше англомовне видання (Dictionary of Modern Written Arabic) – Arabic dictionary for the contemporary written language. Wiesbaden, 1961).

DIN-31635 (1982). Створено Німецьким інститутом стандартизації (DIN – Deutsches Institut für Normung) на основі стандарта DMG.

ISO – Міжнародна асоціація стандартизації (International Organization for Standardization) – розробила низку стандартів використання латиниці для відображення символів інших систем письма – ISO/R 233:1961, ISO 233:1984, ISO 233-2:1993, ISO 3166-2:1998.

BS 4280 (1968). Прийнято Британською організацією стандартів (BS – British Standards Institution).

SAS (поч. XIX ст.) – система транслітерації, що склалася в іспаномовних арабістичних працях (SAS – Spanish Arabist's School).

SATTS – стандартна система арабської технічної транслітерації (Standard Arabic Technical Transliteration System).

UNGEGN (1972). Розроблено групою експертів ООН у географічних назвах (UNGEGN – The United Nations Group of Experts on Geographical Names).

Qalam (1985) – система, орієнтована на точне передання написання.

ArabTeX (з 1992, є близькою до ISO/R 233 і DIN 31635), входить до пакетів TeX і LaTeX.

BT – система транслітерації Букуолтера (Buckwalter Transliteration), розроблена Тімом Букуолтером (Timothy Buckwalter) для компанії Xerox 1990 р.

BATR (Bikdash Arabic Transliteration Rules) є компромісом між Qalam і BT.

Сімейство систем NewWay (з 2006 р.) – ALFB (ALFB NewWay), SIMPLINE.

SEHL (2010) – система, розроблена для носіїв мов, що використовують латиницю.

ALA-LC (1997) – стандарт, прийнятий Американською бібліотечною асоціацією і Бібліотекою Конгресу (ALA-LC American Library Association & Library of Congress).

AlMurih – проста, але самодостатня система, що не використовує додаткових символів.

BGN/PCGN (1956) – система транслітерації арабських географічних назв, пов'язана з домовленостями з латинської транслітерації, прийнятими Американським комітетом з географічних назв (BGN – United States Board on Geographic Names) і Постійним комітетом з географічних назв для офіційного вживання у Великобританії (PCGN – Permanent Committee on Geographical Names for British Official Use).

ACGN (1971). Бейрутську систему транслітерації арабських власних імен було прийнято на Третій арабській конференції з географічних імен (ACGN – Arab Conference on Geographic Names) у 1971 р. Вдосконалена Бейрутська система з'явилася в 1972 р.

I.G.N. (1973) – стандарт, розроблений Французьким національним географічним інститутом (I.G.N. – Institut Géographique National de France). Його іноді називають другим варіантом вдосконаленої Бейрутської системи.

RJGC – Йорданський королівський географічний центр (Royal Jordanian Geographic Centre) у 1981 році прийняв систему транслітерації на основі вдосконаленої Бейрутської системи. Практично збігається з системою BGN/PCGN, за винятком того, що тут літера ё передається через dh, а не через z, а ё передається через a, а не через h.

SES (1930–1972) – Єгипетська система транслітерації ((Arabic) Survey of Egypt System).

SM – система, що склалася внаслідок вживання символів латинської графіки, які використовуються у сучасних мовах, для передання арабської графіки (SM – (Roman) Simplified Modern).

EI – система транслітерації, прийнята в Енциклопедії ісламу (Encyclopaedia of Islam, перше видання Brill, 1913–1938 pp.).

Окрім вищезазначених, існують інші системи транслітерації, наприклад Марокканська (1932) та Ліванська (1963) національні системи. Окремо варто згадати практику транслітерації, яка склалася у різних галузях використання арабської мови поза межами будь-яких стандартів, наприклад транслітерація, що використовується в арабомовних SMS-повідомленнях і чаті, яка не є системою в повному розумінні цього слова й виникла в процесі спілкування в онлайновому режимі (див. табл. 1).

Подібною є ситуація з кириличною транслітерацією. Фонетичний алфавіт на кириличній основі був наведений вже в “лінгвістичній абетці” В.В. Радлова, В.П. Васильєва і Г.К. Залемана РЛА ((Русская лингвистическая азбука) – фонетична транскрипція Російської академії наук на основі російського алфавіту, яка дісталася широке застосування в дореволюційній Росії). Варіанти цього алфавіту зустрічаються у працях вітчизняних арабістів і сходознавців початку ХХ ст. В РЛА для позначення звуків, відсутніх у російській мові, використовуються додаткові символи й спеціальна діакритика. Пізніше акад. І.Ю. Крачковський вирішив дещо вдосконалити російську транскрипцію арабомов-

них текстів і створив дві системи кириличної транскрипції – точну й приблизну. Обидві ці системи використовують російські літери й мають лише один диграф – дж. У точній системі для позначення арабських звуків, відсутніх у російській мові, використовується діакритика, що є близькою до стандартів ISO, у приблизній – використовуються майже винятково російські літери (крім апострофа для «ї» перевернутого апострофа для «ъ») [Сердюченко 1967, 112, 113]. Обидві ці системи знайшли використання у російськомовних публікаціях, на їхній основі базується сучасна кирилична транслітерація арабських текстів [Сердюченко 1967, 112, 113]. Подібно виглядає й українська транслітерація на кириличній основі (див. табл. 1).

Як можна побачити, варіанти передання арабської графіки й позначення дискретних арабських фонем на письмі є доволі численними. Іншими словами, одиниці сегментного рівня представлені у системах транскрипції й транслітерації арабомовних текстів достатньою кількістю варіантів, які дозволяють здійснювати цю транскрипцію та транслітерацію адекватно до їхнього практичного призначення.

Зовсім іншою є ситуація з одиницями супрасегментного рівня. Лише деякі системи транскрипції, переважно фонетичної, приділяють увагу цим одиницям. Наприклад, МФА містить набір спеціальних символів для відображення елементів просодичної структури. Проте ці символи не є уніфікованими й використовуються в досить невеликій кількості публікацій. Утім, у більшості випадків засоби позначення одиниць супрасегментного рівня мають оказіональний характер.

Тим цікавішим є досвід позначення на письмі одиниць супрасегментного рівня АЛМ, що склався у межах різних підходів у різні часи. Наведемо два приклади.

В арабській лінгвістичній традиції з VIII ст. н.е. склалася теорія віршування, засновником якої вважається Аль Халіль бен Ахмад Аль Фарагід (718/19–781). У своїй праці під назвою “كتاب العروض” (“Книга віршування”), яка, на жаль, не дійшла до наших днів, учений подає класифікацію арабських віршованих розмірів. Ця система базується на розрізненні специфічних чергувань довгих та коротких складів, або, на думку класичних арабських філологів, “огласованих та неогласованих літер” [Долинина 1983, 16]. Сам Аль Халіль порівнює віршований рядок з наметом (бейт), що має два схили (піввірші). Піввірш складається зі стоп, кожну з яких можна уявити як вищезгадане сполучення “літер”, представлених конструктивними елементами намету, а саме “кілочками” й “зав’язками”. Чез рівність серед арабів у часи Аль Халіля чіткого визначення складу таке представлення виявилося практичним і знайшло широке використання серед арабських філологів до сьогодні. Залежно від того, які склади утворюють стопу, її елементи можуть бути “зав’язками” й містити один довгий (легка зав’язка) чи два коротких склади (важка зав’язка) або “кілочками” – короткий, потім довгий (прив’язаний кілочок) або довгий, потім короткий склади (відв’язаний кілочок) [Долинина 1983, 17].

У сучасних працях, присвячених арабській фонетиці, також можна зустріти символи на позначення “легких” (типу CV) та “важких” (що містять послідовності типу VCC, V:C, VVC) складів – s і \$ відповідно, – а символ s використовується на позначення складів обох типів, якщо цей склад не впливає на розподіл наголосу [Canepari 2005, 327]. Обидва способи позначення одиниць супрасегментного рівня, хоча й з різних причин, дозволяють абстрагуватися від конкретної фонематичної реалізації складів та їхніх сполучень, а за необхідності – й від окремих характеристик просодичного рівня, таких, як словесний наголос. Як можна побачити, таке представлення структури складів нагадує просодичні шаблони, що знайшли використання у сучасних лінгвістичних дослідженнях.

## ЛІТЕРАТУРА

- Долинина А.А. Из древней арабской поэзии и прозы / Пер. с араб. А.А. Долининой и Вл. В. Полосина. Москва, 1983.
- Зиндер Л.Р. Общая фонетика / Зиндер Л.Р. Москва, 1979.
- Кодзасов С.В., Кривнова О.Ф. Общая фонетика. Москва, 2001.
- Рыбалкин В.С. Арабская лингвистическая традиция: Истоки, творцы, концепции. Киев, 2000.
- Сердюченко Г.П. Русская транскрипция для языков зарубежного Востока. Москва, 1967.
- Canepari L.A. Handbook of Pronunciation / L. Canepari. Lincom Europa, 2005.
- Handbook of the International Phonetic Alphabet. Cambridge University Press, 2005.

Таблиця 1.

Графічний знак	Unicode	SATTS	UNGEGN	DIN 31635	Qalam	SAS	SM	Buckwalter	IPA	BATR	ArabTeX	ALFB	SIMPLine	SEHL	EI	Arabizi	Arabesh + chat	Кирилична транслітерація	Латинська транслітерація
ا	0637	U	۰	۱	T	۱	I	T	t̄	T	۱	θ	۸	۴	t	۶	t, T, TH, ۶	۱	۱
ب	0638	Y	۲	۲	Z	۲	۲	۲	۲	Z	۲	۲	۸	۴	۷	۶	۲, ۶	۳	۲
ت	0639	*	*	*	۴	*	*	*	*	E	*	*	۰	۰	۳	۳	*	*	*
ث	063A	G	گ	گ	ḡ	گ	گ	گ	گ	گ	گ	گ	۸	۶	۳'	۳'	۸	۸	
ف	0641	F	f	f	f	f	f	f	f	f	f	f	f	f	f	f, ph	۲, ۲, ۸,	۹, ۹, ۸	
ق	0642	Q	ڧ	ڧ	q	ڧ	ڧ	ڧ	ڧ	ڧ	ڧ	ڧ	۹	۹	۱	۱	۱	۱	۱
ک	0643	K	k	k	k	k	k	k	k	k	k	k	۸	۸	۲	۲	۸,	۸	
ل	0644	L	ل	ل	l	ل	ل	ل	ل	ل	ل	ل	۱	۱	۱	۱	۱	۱	
م	0645	M	m	m	m	m	m	m	m	m	m	m	۱	۱	۱	۱	۱	۱	
ن	0646	N	n	n	n	n	n	n	n	n	n	n	۱	۱	۱	۱	۱	۱	
ه	0647	H	ه	ه	h	ه	ه	ه	ه	ه	ه	ه	۱	۱	۱	۱	۱	۱	
و	0648	W	w	w	w	w	w	w	w	w	w	w	۰	۰	۰	۰	۰	۰	
ي	064A	Y	y	y	y	y	y	y	y	y	y	y	۱	۱	۱	۱	۱	۱	
ؠ	0629	@	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	۰	۰	۰	۰	۰	۰	
ؠ	0649	I	y	ؠ	ae	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	ؠ	۱	۱	۱	۱	۱	۱	