OBLICZA KOMUNIKACJI 12/2020
https://doi.org/10.19195/2083-5345.12.5

ANNA MAJEWSKA-TWOREK
ORCID: 0000-0003-3997-5991

Uniwersytet Wrocławski
Wydział Filologiczny, Instytut Filologii Polskiej

MONIKA ZAŚKO-ZIELIŃSKA
ORCID: 0000-0001-6333-660X

Uniwersytet Wrocławski
Wydział Filologiczny, Instytut Filologii Polskiej

Funkcjonalny opis transkrypcji jako wstępnego etapu analizy danych mówionych

Abstrakt

Artykuł dotyczy istotnego zagadnienia w badaniach mówionej odmiany języka — sposobu transkrybowania danych mówionych. W tekście przedstawiono główne rodzaje transkrypcji, które zostały zastosowane w różnych polskich i międzynarodowych projektach badawczych. Opisano także zależność między celem badań a typem transkrypcji nagranych wypowiedzi. Podkreślono, że decyzja o sposobie transkrybowania powinna zostać podjęta przez badacza świadomie, by mógł dobrze zrealizować cel swoich poszukiwań naukowych. W końcowej części przedstawiono propozycję typologii transkrypcji danych mówionych. Większość wykorzystanych w artykule przykładów zapisu wypowiedzi ustnej pochodzi z archiwum Pracowni Polszczyzny Mówionej Uniwersytetu Wrocławskiego.

Słowa kluczowe: analiza danych mówionych, transkrypcja tekstu mówionego, fonetyka, dialektologia, analiza konwersacyjna, leksykografia, typologia transkrypcji.

Wstęp

Badanie mówionej odmiany języka jest z jednej strony fascynujące i inspirujące, a z drugiej — bardzo czasochłonne, czasami wręcz żmudne. Między innymi z tego powodu nie każdy chce się analizą i opisywaniem mówionych tekstów zajmować. Przez lata ulotność wypowiedzi mówionej była istotnym czynnikiem hamującym rozwój metod służących zgłębianiu oralności. Jeszcze współcześnie można zauważyć wyjątkową pokusę spoglądania na wypowiedź mówioną z perspektywy zasad, któ-

© for this edition by CNS

rymi rządzi się tekst pisany. Zazwyczaj prowadzi to do schematycznych wniosków o spontanicznej wypowiedzi mówionej: jej niestaranności, nasyceniu błędami czy chociażby — przejęzyczeniami i anakolutami. Trzeba wyraźnie podkreślić, że stosowanie klasycznej teorii składni do opisywania jednostek wypowiedzi mówionej prowadzi na manowce, ponieważ jest ona opracowana na podstawie tekstów literatury pięknej. Spontaniczna wypowiedź mówiona (czyli taka, która jest planowana i realizowana przez mówcę na bieżąco, nie jest wyuczona na pamięć) powstaje w zupełnie inny sposób. Z tego powodu wymaga odrębnych metod i narzędzi analitycznych.

Zanim jednak badacz zadecyduje o metodzie analizy swojego materiału mówionego, musi go we właściwy sposób zgromadzić (czyli potrzebne wypowiedzi po prostu nagrać; por. m.in. Zaśko-Zielińska, Majewska-Tworek, Śleziak, Tworek 2020), a następnie przygotować do właściwego opisu — musi je przetranskrybować. Transkrypcja jest często tak trudnym etapem pracy nad wypowiedzią mówioną, że początkujący badacz jest gotów w zdecydowany sposób ją odrzucić, a nawet — całkowicie zmienić przedmiot swoich zainteresowań. Odrzucają ją także osoby, które z racji swojej profesji nigdy nie powinny tego czynić, na przykład logopedzi.

Niniejszy artykuł ma na celu przedstawienie różnych sposobów zapisu wypowiedzi mówionej oraz wskazanie, w jakich badaniach powinny zostać zastosowane, aby powstałe transkrypty mogły pełnić swą funkcję jak najlepiej. Wniknięcie w zagadnienie transkrybowania prowadzi do jednoznacznego wniosku: typów transkrypcji jest bardzo wiele; zależą one od wielorakich czynników. Wybór sposobu zapisu jest zależny między innymi od celu badawczego lub utylitarnego: czemu przeniesione na pismo (przetranskrybowane) wypowiedzi mają służyć. Innymi słowy jakiemu celowi czynność transkrybenta jest podporządkowana. Świadomość celu jest tu niezmiernie istotna. Narzuca ona nie tylko typ transkrypcji, lecz także szczegółowe zasady, które w ramach danego typu są zazwyczaj modyfikowane. Czyni się to zwykle po to, by w odpowiedni sposób dostosować konwencję transkrypcyjną do szczegółowych celów całego przedsięwzięcia. Ważne jest również, by osoba zajmująca się badaniem mówionego wariantu języka poszczególne typy transkrypcji pojmowała, umiała je nazwać oraz od siebie odróżnić.

Kiedy jest potrzebny zapis danych mówionych?

Zapis analizowanego materiału jest oczywiście niezmiernie ważnym etapem w realizacji każdego projektu zgłębiającego cechy mówionej odmiany języka. Zarówno w badaniach dawnych, jak i współczesnych zapis ten jest bardzo zróżnicowany. Tradycyjnie kojarzy się przede wszystkim z fonetyką (mniej z fonologią) i jej celami badawczymi. Efekty badań fonetycznych przedstawiane głównie w postaci zapisów transkrypcyjnych — poza samą fonetyką rozumianą tutaj jako dyscyplina badań podstawowych — znajdują się w pracach z zakresu dialektologii, językoznawstwa normatywnego, leksykografii (w słownikach różnego typu) czy logopedii.

© for this edition by CNS

Przykłady zapisywane alfabetem transkrypcyjnym różnią się zarówno stopniem szczegółowości, jak i konwencją przyjętą przez autora. W badaniach slawistycznych (a więc i polonistycznych) największą zmianą, którą obserwuje się od kilku lat w stosowaniu transkrypcji, jest coraz powszechniejsze wykorzystywanie alfabetu międzynarodowego IPA. Argument, że alfabet slawistyczny lepiej oddaje cechy fonetyczne języków słowiańskich, odchodzi do przeszłości. Transkrypcję zupełnie innego typu stosują badacze wykorzystujący metodę analizy konwersacyjnej. Sięgają po nią przy zgłębianiu innych cech wypowiedzi monologowych czy dialogowych niż wyłącznie warstwa fonetyczna; w takich przypadkach nie ma potrzeby stosowania znaków z transkrypcji fonetycznej. Istotniejsze są bowiem zupełnie inne znaczniki, które są do pewnego stopnia stałe, mocno już skonwencjonalizowane, ale też każdorazowo modyfikowane na potrzeby określonego projektu.

O wybranych typach i konwencjach transkrypcyjnych wraz z przykładami ich zastosowania w badaniach

W dalszej części pracy — z konieczności w dużym skrócie — zostaną omówione główne cechy zapisu transkrypcyjnego w wymienionych wcześniej typach badań, a także wybrane trudności towarzyszące niełatwej sztuce notowania tekstów mówionych. Część przykładów będzie pochodziła z wybranych, zrealizowanych już projektów, przeprowadzonych w ramach Pracowni Polszczyzny Mówionej Uniwersytetu Wrocławskiego.

1. Zapisywanie wymowy w badaniach fonetycznych

Zapisywanie wymowy jest umiejętnością niezbędną dla osób zajmujących się badaniami fonetycznymi różnego typu w obrębie fonetyki jako nauki podstawowej, a także nauk stosowanych, między innymi dialektologii czy logopedii. Podstawowe wzory zapisów polskiej wymowy można znaleźć w pracach dotyczących bezpośrednio fonetyki współczesnego języka polskiego. Zwykle podstawowe zestawienie znaków znajduje się w podręcznikach w postaci list — często uporządkowanych tabelarycznie w dwóch konwencjach jednocześnie: slawistycznej i międzynarodowej (Dukiewicz 1995: 15). Pogłębione badania fonetyczne często wymagają wprowadzania dodatkowych znaków diakrytycznych, które są ustalane w osobnych publikacjach (Porayski-Pomsta, Kwasiborska-Dudek, Jauer-Niworowska, Lipiec, Więcek-Poborczyk, Golanowska, Malinowska 2013) lub — jeśli jest taka konieczność — samodzielnie proponowane przez autorów w trakcie realizowania indywidualnego projektu badawczego. Jak już wspomniano, we współczesnych pracach slawistycznych powoli za oczywiste uznaje się stosowanie alfabetu międzynarodowego,

© for this edition by CNS

co należy uznać za istotny postęp w upowszechnianiu wyników badań w obiegu naukowym, choć jeszcze pięć lat temu Piotr Rybka (2015: 5) słusznie napisał:

> Międzynarodowy Alfabet Fonetyczny (IPA) jest powszechnie znanym systemem transkrypcji stosowanym w różnych dziedzinach badań nad mową, lecz bardzo rzadko używanym w obrębie językoznawstwa polskiego, nie tylko w badaniach nad językiem polskim, ale także nad innymi językami słowiańskimi.

Poza zasygnalizowanymi już typami transkrypcji należy wyróżnić transkrypcję fonetyczną od fonologicznej. Transkrypcja fonetyczna używana jest w dwu odmianach, określanych w anglojęzycznych publikacjach jako *broad* albo *phonemic transcription* i *narrow* albo *allophonic transcription*. W transkrypcji fonologicznej (fonemicznej, 'broad') jeden znak służy do zapisu całej klasy głosek reprezentujących jeden fonem, np. [**r**ɑk// rtɛɲt͡ɕ], a w zapisie fonetycznym dla każdej głoski istnieje odrębny symbol, np. [rɑk/ Rak// rtɛɲt͡ɕ/ r̩tɛɲt͡ɕ]. Każda z nich znajduje swoje zastosowanie w określonym typie badań. Transkrypcja fonetyczna służy notowaniu najbardziej subtelnych różnic między dźwiękami mowy i wymaga doskonałego słuchu fonetycznego, czyli takiego, który wykracza poza zdolności audytywne przeciętnego użytkownika polszczyzny. Przeciętny użytkownik języka na co dzień posługuje się słuchem fonologicznym, który kształtuje się w procesie ontogenezy mowy. Niezbędnym warunkiem jego ukształtowania się jest prawidłowy słuch fizyczny oraz naturalne zanurzenie w odpowiednio bogatym językowo środowisku. Posługiwanie się słuchem fonologicznym jest wystarczające do prawidłowego posługiwania się językiem na co dzień. Zwykłemu użytkownikowi języka nie przysporzy kłopotu słuchowe odróżnienie głosek [r] od [l]. Natomiast może mieć (choć — oczywiście — nie musi) trudności w odróżnieniu od siebie głosek, na przykład [r/ R/ r̩], ponieważ występujące między nimi różnice nie decydują o różnicy znaczenia wyrazów, w których występują: realizacja głoskowa znaczona symbolem [R] będzie charakteryzowana jako wada wymowy, a dźwięk oznaczany symbolem [r̩] pojawi się ewentualnie w określonym otoczeniu — jest więc zależny od kontekstu fonetycznego. Posługiwanie się alfabetem fonetycznym wymaga zatem treningu słuchowego i dużej wrażliwości na niedystynktywne różnice między poszczególnymi dźwiękami mowy. Efektem takiego treningu może być bardzo szczegółowa analiza (por. przykład 1). W przykładzie 1 autorka transkrypcji nie notowała pauz, ponieważ nie były one przedmiotem zainteresowania badawczego. Tekst był czytany na głos bez wcześniejszego przygotowania. Można by więc orzec, że jest to wypowiedź pierwotnie pisana, a mówiona — wtórnie.

Ponadto Anita Lorenc (2016) wyróżniła w ramach transkrypcji fonetycznej wąskiej *transkrypcję impresjonistyczną* i *systematyczną*. Pierwsza odnosi się do zapisu wszystkich stwierdzonych cech wymowy bardzo mocno zniekształconej, w której analizie — szczególnie na jej początkowym etapie — nie sposób odnieść się do określonych miejsc systemu fonologicznego. Stosuje się w niej wiele znaków diakrytycznych, proponowanych na przykład w pracy Porayskiego-Pomsty i innych (2013). Natomiast *systematyczna transkrypcja* to inaczej zapis allofoniczny, bliższy badaczom fonetyki ogólnej, regionalnej czy dialektologicznej.

© for this edition by CNS

Przykład 1. Transkrypcja w IPA tekstu czytanego po polsku przez osobę rosyjskojęzyczną (Gaudy 2017)

ɲigdɨ̞ ɲɛbɨ̞wam zbʷɨt tɔvarʲisce dlatɛgɔ ɲigdɨ̞ ɲɛm͡ʲawɛm sm͡ʲal̪ʷɔstʲ͡i dɔ d͡zʲɛf t͡ɕɔt vmɔjə̞m studɛnt͡scm ʒɨ̞t͡sʲu zat͡ɕɛwa mʲijɛdnak dɔsk͡vʲɛrt͡satɕ samʷɔtnɔst͡sʲ jagbɨ vɔtpɔv͡ʲɛd͡zʲi namɔjɛ mɔdʲitvʷɨ ft͡ɕɔraj v͡ʲɛt͡ɕɔrɛm zad͡zvɔɲiwa dɔmɲɛ t͡ʃʲɔt͡ʃʲa spʲ͡etrkɔva ɲɛ zdawawɛm sɔbʲ͡e sprav͡ʷɨ stɛgɔ jagbard͡zə̞ tɛn tə̞l̪ʷɛfɔn ɔdm͡ʲɛɲi mojə̞ ʒɨ̞t͡sʲɛ t͡ɕɔt͡ɕa xɛlɛna pɔprɔsʲiwamɲɛ ʒɛbə̞m spɔtkawɕɛ d͡zʲis saʃɔm saʃa pɔmɨʃʲ͡awɛm jacɛ pʲ͡ɛɲkɲɛ imʲ͡ɛ ʃɨ̞pkɔ pɔbʲ͡ɛgwɛm pɔkv͡ʲatɨ̞ pɔt͡ʃʲim udawɛmsʲ͡ɛ nam͡ʲɛsʲcɛ spɔtkaɲa t͡ɕɛkawɛm abgrɨzajɔns paznɔkt͡ʃʲe iviɔbraʒajə̞ns sɔbʲ͡ɛ spaɲawə̞m dwugɔnɔgɔm bl̪ʷɔndincɛ ɔɲebʲ͡ɛscx ɔt͡ɕax ɲedal̪ʷə̞ɕə̞ ukrɨt͡ʃʲ mɔjɛgɔ zd͡zʲiv͡ʲɛɲa cɛdɨ saʃa ɔkazal̪ʷɕə̞ mɔjem dalɛcm kuzʲinɛm t͡suɕ m͡ʲawə̞m zrɔbʲit͡ʃʲ pɔxvɨʲʲi ɲɛzrɛt͡ʃʲnɛj tʲiʃɨ vrɛnt͡ʃʲiwɛm mu bucɛt ipɔʃʲʲiʃmɨ razɛm napʲivɔ

Innym typem zapisu wypowiedzi mówionej jest transkrypcja półfonetyczna, czyli uproszczona transkrypcja fonetyczna, zaproponowana w latach pięćdziesiątych XX wieku przez Alfreda Zarębę (1955). Stosuje się ją często w badaniach dialektologicznych (zarówno w analizach gwar wiejskich, jak i miejskich). W takim zapisie przeważa zapis ortograficzny, natomiast miejsca szczególnie charakterystyczne dla danej gwary oznaczane są w specjalny sposób za pomocą wcześniej ustalonych znaków przez badacza lub znaków powszechnie stosowanych w danym typie analizy.

Przykład 2. Transkrypcja półfonetyczna gwarowej wypowiedzi kobiety z Łososiny Górnej (www1)

> Krowy ch^{ł}owało sie p^{ł}o to, żeby mieć mlyk^{ł}o dla ^{o}dzieci, z mlyka zrobić syra i masła, a to, co zostawało, ta maślonka ^{j}i se^{y}rwotka, to robiło sie nawet zo^{u}r do zie^{y}mnioków na ^{ł}obiod… albo p^{ł}o prostu p^{ł}opijało sie tom maślonkom takie ^{ł}okrasone zie^{y}mnioki[1].

Samodzielne rozszerzanie systemu zapisu cech gwarowych przez poszczególnych badaczy (często także amatorów i zwykłych pasjonatów gwar) spowodowało niemały chaos, na co zwrócili uwagę autorzy książki *Standard zapisu fonetycznego polskich cech gwarowych* (Czesak, Okoniowa, Waniakowa 2004). Współcześnie dąży się więc do ujednolicenia systemu zapisywania fonetycznego tekstów gwarowych, czemu ma służyć wspomniana publikacja. Jej autorzy słusznie zwracają uwagę,

[1] Trzeba podkreślić, że oprócz cytowanego fragmentu na stronie portalu dialektologicznego został umieszczony plik dźwiękowy z nagraniem (wraz z podstawowymi metadanymi mówczyni), co umożliwia samodzielny odsłuch i interpretację danych. Zazwyczaj jednak nagrania terenowe na zawsze zostały oderwane od swego źródła, co współcześnie uniemożliwia ich dalszą eksplorację. Innym problemem jest znikome archiwizowanie danych o nagrywanych mówcach, co już na starcie uniemożliwia analizę zgromadzonych niegdyś danych językowych (o archwizowaniu danych mówionych zob. Zaśko-Zielińska, Majewska-Tworek, Śleziak, Tworek 2020).

© for this edition by CNS

że zapis tekstu jest zawsze interpretacją słuchową, której nie zobiektywizują szczegółowe znaki diakrytyczne. Wydaje się zatem (o czym już jednak autorzy nie piszą), że najlepszym sposobem na rozwiązanie tego problemu jest konsekwentne integrowanie zapisu z plikiem dźwiękowym, aby umożliwić bieżące weryfikowanie audytywnej interpretacji autora transkryptów.

Z problemem zapisu danych mówionych zderzyli się uczestnicy projektu „Polszczyzna mówiona mieszkańców miast" w latach siedemdziesiątych–dziewięćdziesiątych XX wieku. Poszczególne zespoły podejmowały bowiem różne decyzje. Uproszczoną transkrypcję fonetyczną zastosowano w wydaniu *Tekstów języka mówionego mieszkańców miast Górnego Śląska i Zagłębia* (Lubaś 1978: 80). Władysław Lubaś (1978: 351) uzasadnił decyzję zespołu badaczy następująco: „Zastosowanie uproszczonej transkrypcji podyktowane jest liczeniem się z odbiorcami spoza dyscyplin językoznawczych". W wydanych tekstach katowickich zmodyfikowano konwencję A. Zaręby, żeby wyeksponować w transkryptach cechy oralności: „odnotowuje się inne zjawiska charakterystyczne tylko dla mówionej odmiany języka, np. wtrącenia, powtórzenia, zająknięcia, długie samogłoski i spółgłoski, osobliwe akcenty oraz zjawiska parajęzykowe, np. śmiech" (Lubaś 1978: 351–352). Natomiast zespół krakowski zadecydował o zapisaniu nagranych wypowiedzi z zastosowaniem transkrypcji slawistycznej[2]. Jednocześnie w ramach tego samego projektu naukowcy wrocławscy w IFP UWr swoje nagrania spisali alfabetem ortograficznym. Transkrypty były przygotowywane do wydania, jednak nigdy drukiem się nie ukazały. Mają one dziś dużą wartość historyczną i sentymentalną[3]. Zamieszczamy zatem krótki fragment zapisu nagrania zatytułowanego „Zebranie zakładu".

Ważnym obszarem, w którym staranna transkrypcja powinna znaleźć swoje zastosowanie, jest logopedia. Jest ona potrzebna zwłaszcza w przygotowywaniu diagnozy sprawności artykulacyjnej (dziecka czy dorosłego). Anita Lorenc zauważa, że w dążeniu do obiektywizacji zapisu wypowiedzi mówionej autorzy transkrypcji wykorzystują różne dodatkowe dane. W wypadku transkrypcji służącej ocenie logopedycznej takiej obiektywizacji pomaga obserwacja pracy narządów artykulacyjnych:

> W diagnozie logopedycznej transkrypcja wykonywana jest nie tylko w oparciu o słuchową ocenę artykulacji, ale także wzrokową. Obserwacja położenia narządów mowy podczas artykulacji (np. w celu odróżnienia artykulacji apikalnych od laminalnych i określenia położenia masy języka wewnątrz jamy ustnej) służy oczywiście trafności diagnozy, a zarazem obiektywizacji samego zapisu fonetycznego. (Lorenc 2016: 136)

[2] Podano dwa główne powody: „Precyzyjny zapis pozwala dostrzec zróżnicowanie systemu fonetycznego polszczyzny mówionej. Zastosowanie pisowni slawistycznej, a nie międzynarodowej podyktowane było zwłaszcza tym, że teksty przeznaczone są również dla celów dydaktycznych" (Dunaj 1979: 15).

[3] Więcej o odnalezionym zbiorze i jego ponownym opracowaniu zob. Majewska-Tworek, Zaśko-Zielińska, Pęzik 2020.

© for this edition by CNS

Przykład 3. „Zebranie zakładu" — fragment (ze zbioru transkrypów Pracowni Polszczyzny Mówionej z lat osiemdziesiątych i dziewięćdziesiątych XX wieku; sygn. MW_001)

Kam.: Tak Janusz ale tutaj szefowie zaznaczyli, że pod ten temat
Waldy: szerszy on będzie
yy ludzi, którzy mają jakieś swoje
tematy zasadnicze, znaczy tam Waldek o czymś myśli.
Ala: Nie tylko teoretyczne, wiesz
Kam.: Ty będziesz dalej robił swoje, oni swoje, tylko żeby było sformułowane
Waldy: Żeby mieściło się to ...
Kam.: Tak, żeby mieściło się to Waldy pod ogólnym tytułem
Bo trudno żebyśmy znowu grupy ludzi tworzyli, żeby robili...
Anus.: No dobrze, ale od kierownika wymaga się przynajmniej tego, żeby się trochę znał na rzeczy, a ja się nie znam na wszystkim
Sic.: eyyy tu możemy uspokoić kierownika że tu są te dwie rzeczy
Niec: kto zna się na wszystkim

Trzeba podkreślić, że w codziennej pracy logopedy diagnosty zapis fonetyczny jest środkiem, a nie celem samym w sobie. Logopeda powinien sprawnie posługiwać się znakami transkrypcyjnymi do zapisywania pojedynczych przykładów zaburzonej lub poprawnej artykulacji (pojedynczych wyrazów czy fraz zdaniowych, rzadziej całych zdań). Umożliwi mu to dobra znajomość fonetyki języka polskiego oraz trening w zakresie transkrybowania tekstów ze słuchu, a także opracowania z dodatkowych znaków diakrytycznych służących zapisywaniu głosek zdeformowanych (Porayski-Pomsta *et al.* 2013).

2. Transkrypcja w leksykografii

Słownik ogólny jako publikacja, która dokumentuje rozwój języka w określonym momencie, powinien uwzględniać również mówioną odmianę języka. Jako że opracowania tego typu przeznaczone są dla szerokiego kręgu odbiorców, który współtworzą głównie rodzimi użytkownicy języka, zazwyczaj zakłada się, że tylko niektóre hasła muszą być uzupełnione informacją o wymowie, a wykorzystany zapis fonetyczny powinien być zrozumiały dla każdej osoby, nawet jeśli nie ma ona odpowiedniego przygotowania. Redaktorzy słowników najczęściej stosują transkrypcję ortograficzną, którą w różny sposób nazywają, aby podjęta decyzja była zrozumiała dla użytkowników, na przykład:

— „(wymowa) podawana jest za pomocą znaków alfabetu łacińskiego, nieznacznie tylko poszerzonego" (WSJP);

© for this edition by CNS

— „zapis wymowy oparty na zwykłym alfabecie rozszerzonym o dwa dodatkowe znaki" (ISJP 2000: LV);

— „zastosowaliśmy pisownię uproszczoną, półfonetyczną" (PSPWP 1994: 38);

— „wymowa jest zapisywana kursywą z zastosowaniem najprostszych znaków alfabetu polszczyzny literackiej" (SWiONO 1997: 137).

Aby odróżnić zapis wymowy od zapisu ortograficznego w pozostałej części hasła, stosuje się nawiasy kwadratowe lub kursywę. Miejsce akcentu wskazywane jest przez wytłuszczenie lub podkreślenie sylaby i dotyczy formy podstawowej wyrazu hasłowego.

Ze względu na to, że zapis wymowy w słownikach ogólnych czy poprawnościowych nie jest stosowany w każdym haśle, transkrypcja wybranych wyrazów staje się problemem marginalnym, na którego rozwiązanie mają czasami wpływ także czynniki typograficzne (Mróz 2016).

Na międzynarodową transkrypcję fonetyczną zdecydowano się w *Słowniku wymowy polskiej*, który jest opracowaniem bardziej specjalistycznym, mimo że został zaplanowany dla „szerokiego kręgu odbiorców, zwłaszcza dla nauczycieli, studentów, pracowników radia i telewizji, a także dla cudzoziemców" (SWP 1997: VI). We wstępie redaktorzy argumentują swój wybór, jednocześnie zauważając, iż przeciw stosowaniu zapisu międzynarodowego przemawia to, że w polskich słownikach poprawnościowych stosowana jest transkrypcja uproszczona, w pracach z fonetyki artykulacyjnej — transkrypcja slawistyczna, a jedynie w pracach z fonetyki i fonologii akustycznej oraz w transkrypcji innych języków korzysta się z pisowni międzynarodowej, która poza tym jest nieznana (SWP 1997: XXIV). Im węższa i bardziej profesjonalna jest grupa odbiorców słownika, w związku z czym wzrasta też troska o szczegółowość danych, tym większą wagę przykłada się do precyzji transkrypcji. Z kolei *Słownik wariantywności fonetycznej współczesnej polszczyzny* został opracowany na potrzeby lingwistyki, logopedii i glottodydaktyki (SWFWP 2005: 7). Wykorzystano w nim alfabet międzynarodowy IPA, który uzupełniono jeszcze dodatkowymi symbolami (SWFWP 2005: 10). Udokumentowane w słowniku fakultatywne realizacje poszczególnych wyrazów zostały zapisane w transkrypcji wariantywnej, oddającej wymowę różnych mówców. Poprzedza je zapis modelu wymowy, który pozwala na uporządkowanie tych wariantów, a jednocześnie ułatwia przeszukiwanie słownika.

Należy podkreślić, że wybór transkrypcji w słowniku zależy głównie od zaplanowanej grupy użytkowników i typu słownika. Gdy jest on ogólny, w warstwie poświęconej wymowie ma zwykle charakter preskryptywny. Wymaga wtedy tylko o tyle dokładnej transkrypcji, aby była rodzajem wskazówek wymawianiowych dla rodzimych użytkowników języka. Natomiast gdy słownik przeznaczony jest dla specjalistów, staje się opracowaniem precyzyjnie dokumentującym sposób wymowy i do tego celu dostosowuje sposób transkrypcji.

© for this edition by CNS

3. Analiza konwersacyjna

Analiza konwersacyjna to metoda prowadzenia badań komunikacji mówionej, przede wszystkim nieformalnej, na podstawie nagrań audio lub wideo oraz sporządzonych transkryptów. Wywodzi się ona z tradycji socjologii etnometodologicznej (Levinson 2010: 341), a obecnie wykorzystywana jest w ramach różnych dyscyplin, na przykład antropologii, psychologii, socjologii czy językoznawstwa. Tym bardziej odmienne są cele gromadzenia danych tą metodą oraz dostosowane do nich typy transkrypcji, które różnią się od siebie pod względem stopnia szczegółowości. Za każdym razem badacz podejmuje decyzję o sposobie transkrypcji, uwzględniając między innymi sformułowane wcześniej cele, liczbę uczestników czy też jakość nagrania zależną od wykorzystanego sprzętu. W analizie konwersacyjnej nie ma jednego, uniwersalnego sposobu zapisu, ale istnieje kilka, które najczęściej są podstawą transkrypcji lub stają się podstawą modyfikacji, na przykład autorstwa Gail H. Jefferson (2004: 13–31) czy GAT2 (Selting, Auer, Barden 1998). Badacze często stosują konwencję transkrypcyjną opracowaną na potrzeby konkretnej analizy i wtedy w swojej pracy umieszczają informacje o repertuarze użytych znaków lub sięgają po jedną z form zapisu, która pojawiła się wcześniej w innych pracach, a jest adekwatna do bieżących potrzeb analizy.

W procesie transkrypcji zaleca się, aby zacząć od wstępnej, orientacyjnej transkrypcji (*rough transcription*), w której zapisuje się treść werbalną, jej interakcyjne przyporządkowanie, natomiast pomija się symbole między innymi związane z prozodią, gestami czy pauzami. Pozwala ona jednak na wstępny ogląd materiału i podjęcie decyzji, czy szczegółowo transkrybowana będzie całość, czy tylko wybrane fragmenty (Moore, Llompart 2017: 410). Taki podstawowy typ transkrypcji doskonale pełni funkcję dydaktyczną, gdyż może być ćwiczeniem, które pokazuje, jak bardzo różnią się od siebie pisana i mówiona odmiana języka. Studenci czasami są mocno zaskoczeni efektem transkrypcji i wydaje im się, że dzięki niej zauważyli, jak niepoprawnie wypowiada się osoba nagrywana. Ta błędna konkluzja jest idealnym punktem wyjścia do dyskusji na temat cech odmiany mówionej języka.

Drugi etap transkrypcji wiąże się z decyzją odnośnie do wyboru systemu transkrypcji, czyli zestawu symboli, który będzie używany w zapisie. Tutaj także można przyjąć różne stopnie dokładności transkrypcji, co wiąże się z wykorzystaniem mniejszej lub większej liczby symboli w ramach transkrypcji minimalnej (*minimal transcript*), podstawowej (*basic transcript*) czy szczegółowej (*fine transcript*).

Do stałych elementów transkryptu należy jego **rama**, która obejmuje nagłówek z tytułem utrwalonego nagrania oraz najważniejszymi informacjami, na przykład o sposobie transkrypcji, danych transkrybenta, czasie trwania rozmowy. Ponadto do ramy przynależy lewa kolumna z numerowanymi wersami transkrypcji. Ułatwiają one przeszukiwanie transkryptu, a gdy znajduje się on w aneksie, pozwalają na podawanie w analizie odpowiednich numerów wersów, które poszerzają zakres ilustracji opisywanych zjawisk.

© for this edition by CNS

W swojej **zasadniczej części** każdy transkrypt zawiera zapisaną najczęściej ortograficznie treść rozmowy. Czasami nagrany materiał (na przykład tekst gwarowy czy wypowiedź osoby bilingwalnej) może wpłynąć na decyzję o zastosowaniu transkrypcji fonetycznej lub półfonetycznej (por. PSPWP 1994). W zapisie rezygnujemy z wielkich liter w funkcji składniowej; czasami pozostają one tylko jako element nazwy własnej. Nie ma także interpunkcji, a pojawiające się wyjątkowo znaki mają zupełnie inne funkcje. Każda wypowiedziana replika jest przypisana mówcy, który oznaczony jest w odpowiedni sposób, na przykład: K, M, Dziecko1, Dorosły2, Sąsiad, Kobieta. Zapis ten służy anonimizacji danych, która może obejmować również inne informacje zastępowane w transkrypcji zaplanowanymi wcześniej oznaczeniami, na przykład: <nazwisko>, <adres>, <telefon>, <nazwa miejscowa>. Treść rozmowy ujęta jest zatem w transkrypcie w ramach replik przyporządkowanym poszczególnym mówcom, a całość zapisu uzupełniają symbole, które oddają segmentację tekstu, warstwę prozodyczną i informują o kontekście wypowiedzi. Zazwyczaj za punkt wyjścia uznaje się propozycję symboli transkrypcyjnych opracowaną przez Gail H. Jefferson (1984)[4], która obejmuje osiemnaście znaków:

[ ]	nakładanie się replik
=	zetknięcie dwóch replik różnych interlokutorów
(0.4)	pauza w granicach jednej sekundy
(.)	mikropauza, nie dłuższa niż jedna dziesiąta sekundy
. lub ↓	intonacja opadająca
? lub ↑	intonacja wznosząca
,	intonacja zawieszona (*continuing intonation*)
-	urwanie wypowiedzi
>słowo<	wyraźnie szybciej wypowiedziane słowo lub fragment
<słowo>	wyraźnie wolniej wypowiedziane słowo lub fragment
◦	wyraźnie ciszej wypowiedziane słowo lub fragment
SŁOWO	wyraz wypowiedziany wyraźnie głośniej niż pozostała część rozmowy
<u>Sło</u>wo	akcent na słowo, część słowa lub fragment wypowiedzi
sło::wo	przedłużona sylaba lub dźwięk
hhh	słyszalny wydech
.hhh	słyszalny wdech
(słowo)	niezbyt wyraźnie słyszane słowo lub fragment
((kontynuuje))	anotacja niewerbalnej aktywności mówcy.

Rycina 1. System notacji według Gail H. Jefferson (1984)

W zależności od celu badań liczba symboli może się zwiększać. Część z nich pochodzi z podstawowego zestawu, ale dzięki zwielokrotnieniu może wskazywać na zwiększenie natężenia cechy, na przykład: ↓ ↓↓ : :: ::: h hh hhh; inne mogą być kombinacją znaków wyjściowych, na przykład ◦hh. Czasami wprowadzane są dodatkowe symbole, które zazwyczaj można zaczerpnąć z poszerzonych zestawów notacji, które wcześniej zostały użyte przez innych badaczy. Do uzupełniania symboli mogą skłaniać zjawiska dostrzeżone w materiale, które reprezentują dane zjawisko,

[4] Na gruncie polskim system ten został spopularyzowany w pracy Doroty Rancew-Sikory (zob. np. Rancew-Sikora 2007).

© for this edition by CNS

ale w nagraniu pojawiają się w pewnym zróżnicowaniu, a ich opis jest istotny w analizie. Czasami odnotowuje się na przykład różne rodzaje śmiechu, gdy mówca:

— w idiolektalny sposób realizuje śmiech: haha, hehe, hihi;
— śmieje się w przerwie między słowami ((śmiech));
— mówi ze śmiechem <<śmiejąc się>*policja*>;
— mówi śmiejącym głosem <<:-)>*więc*> (Rogoziński, Zarzeczny 2012).

W sytuacji analizy wkładu konwersacyjnego w polilogu istotny może być też stopień głośności wypowiedzi (Selting, Auer, Barden 1998: 39), na przykład:

— <<f>*słowo*> *forte*, *loud*;
— <<ff>*słowo*> *fortissimo*, *very loud*;
— <<p>*słowo*> *piano*, *soft*;
— <<pp>*słowo*> *pianissimo*, *very soft*.

Bardzo istotne są także uszczegółowienia dotyczące pauz, które pozwalają na dokładne odnotowanie czasu ich trwania, gdy transkrypt nie jest sporządzany za pomocą programu komputerowego, oraz opis sposobu realizacji pauz wypełnionych, na przykład *yy*, *ee*, *hm*, *no*.

Mimo że do utrwalenia treści rozmów w analizie konwersacyjnej oprócz nagrania używamy zapisu, nie można zapominać, że transkrypt nie jest podporządkowany regułom pisanej odmiany języka. Rządzą nim zupełnie inne zasady, które dobieramy w odniesieniu do celu badań, eksponując symbolami między innymi warstwę prozodyczną tekstu lub jego segmentację.

Przykład 4. Komentarz sportowy w bezpośrednim przekazie meczu piłki nożnej (Czechowski 2020)

no = hhh< (.) chciał zaskoczyć joeharta (.) ale to nie taka prosta sprawa … obraniak lewandowski dobre przyjęcie (.) czeka (.) > na wolne pole do grosickiego grosicki pole karne GROSICKI AH::! < z pierszej piłki (0.7) no i teraz zabrakło troszkę (.) zimnej krwi (.) stwarzamy w tym meczu sytuacje nieporównywalnie więcej tych sytuacji niż anglicy (0.6) ci (1) no (.) dotknęli nas (krótki śmiech) albo .hh użądlili (0.7) zsz=ze stałego fragmentu, … polacy: (.) może to jest ten moment GLI:::K GLIK GLIK GLIK!

Przykład 5. Rozmowa z dzieckiem bilingwalnym (Hordejuk 2020)

142 B: om y takie dwa ludzie ym y m:ał- mają taki dom na y:: (.) °o nie°. (.) treehouse! I
143 mają tak y: sześćdziesiąt pięć tak (.) stories. i ym y:: y jest o kiedy jeden z i- z nich y:
144 zrobił taki c- ym y: ((mlaśnięcie)) y rzecz co może z- zabrać y cię gdzieś inny- innie- y:
145 tak. y:: gdzieś innego. i potem .hh ł- ym używali to y. i poszli do y miejsca l- jak egipt i:
146 wtedy dinozaury były, i pot- i- i- musi- musiły takie problemy y: °zrobić°? (3) fix
oblems.

4. Zapis wypowiedzi mówionej w programie cyfrowym jako metoda multifunkcjonalna

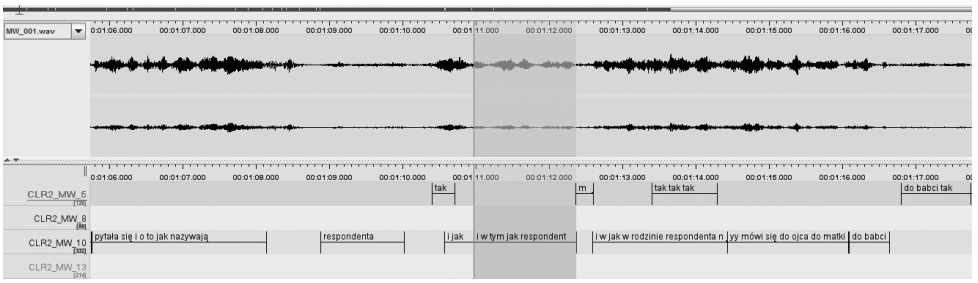
Współcześnie można korzystać z jeszcze innego sposobu zapisu danych mówionych. Służą temu specjalne programy, na przykład ELAN czy EXMARaLDA. Dwa z nich były już wykorzystywane przez Pracownię Polszczyzny Mówionej w opracowywa-

© for this edition by CNS

niu danych językowych (w projektach „Mowa_Wrocław" i „Biografie językowe Dolnoślązaków — pokolenie najstarsze"). Mają one wielką przewagę nad omówionymi technikami zapisu. Wielką zaletą transkrybowania w programie cyfrowym jest bowiem ścisły związek pliku dźwiękowego z linią czasu (określającą czas trwania analizowanej wypowiedzi) i linią zapisu tekstu mówionego. Umożliwia to analizę danych mówionych z dokładnością do setnej sekundy, a więc sprzyja zauważaniu i notowaniu wszystkich zjawisk towarzyszących mówieniu: pauz różnego typu, autokorekt, symptomów retardacji czy cech wynikających z interakcji werbalnej interlokutorów. Transkrypcji i odsłuchowi towarzyszy nieskomplikowany oscylogram. Niezmiernie ułatwia on wstępną analizę już na etapie zapisu danych mówionych. Opracowane dane, anotowane zgodnie z założeniami konkretnego projektu, można transportować do następnych programów, które pomogą opracować kolejne istotne w danym badaniu wskaźniki (na przykład cechy prozodyczne czy ścisłe informacje morfosyntaktyczne).

Przykład 6 ilustruje zapis wypowiedzi w programie ELAN — jednym z najczęściej i najchętniej wykorzystywanych programów tego typu. Ilustruje go fragment materiału językowego pochodzącego z nagrania „Zebranie zakładu". Jest to nagranie dyskusji wrocławskich językoznawców przystępujących do badań nad polszczyzną mówioną wrocławian w latach dziewięćdziesiątych XX wieku. Brało w niej udział kilkunastu naukowców. Każdemu została przyporządkowana jedna linia zapisu (linia werbalna). Na prezentowanym zrzucie widać linie czterech mówców i fragment dialogu dwojga z nich.

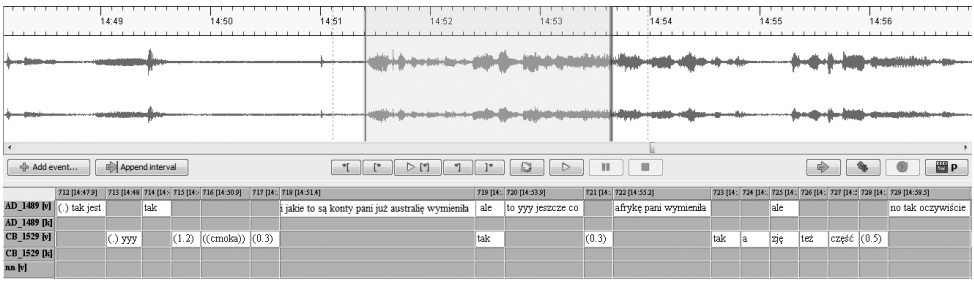
Przykład 6. Partytura programu ELAN — fragment zapisu (zrzut z ekranu)



Z kolei przykład 7 przedstawia fragment zapisu rozmowy (egzaminacyjnej) w programie EXMARaLDA. Praktyka dowodzi, że jest to transkrypcja sprzyjająca badaniu interakcji werbalnej między rozmówcami. Bardzo dobrze widać bowiem nakładanie się wypowiedzi współrozmówców, wzajemne korygowanie się czy przerywanie sobie nawzajem czy wpadanie sobie w słowo. Mając do dyspozycji odpowiednio zrównoważony materiał, można prowadzić interesujące i istotne badania tego właśnie typu.

© for this edition by CNS

Przykład 7. Partytura programu EXMARaLDA — fragment zapisu (zrzut z ekranu)



Transkrybowanie w programach cyfrowych jest niezmiernie czaso- i kosztochłonne. Jednak dzięki niemu powstaje archiwum, które wielokrotnie może służyć w różnych projektach naukowych. Może przez długie lata być wykorzystywane jako rzetelny materiał źródłowy oraz porównawczy.

Transkrypcje — propozycja typologii według różnych kryteriów

Przedstawione typy transkrypcji to tylko niektóre możliwości utrwalania tekstu mówionego za pomocą pisma. Różnią się doborem znaków, stopniem szczegółowości, zakresem wykorzystania technologii do ich tworzenia oraz założonym przez transkrybentów celem badań. Przedstawiamy zatem propozycję typologii transkrypcji, która pozwala wstępnie zapoznać się z ich dużym zróżnicowaniem, świadomie dopasować typ zapisu do ustalonych potrzeb badawczych, a także — co bardzo istotne — konsekwentnie stosować wybrany typ transkrypcji w ramach prowadzonych badań.

1. Ze względu na różnice stosowanego kodu w transkrypcji:
   1.1. stenografia,
   1.2. transliteracja,
   1.3. ortograficzna,
   1.4. ortograficzna z pojedynczymi znakami fonetycznymi (słowniki),
   1.5. ortograficzna z wybranymi znakami specjalnymi (na przykład z zakresu prozodii, elementów niewerbalnych),
   1.6. fonematyczna (ortograficzna na przykład z zapisem realizacji nosowości),
   1.7. fonetyczna (tu różne alfabety),
   1.8. fonologiczna (szeroka fonetyczna),
   1.9. z tagowaniem/anotacją (warstwy fonetycznej, leksykalnej, słowotwórczej, syntaktycznej).
2. Ze względu na kompletność i szczegółowość:
   2.1. całkowita (cała wypowiedź),
   2.2. częściowa (fragmenty),
   2.3. tylko warstwa werbalna,
   2.4. z ingerencją osoby trzeciej (na przykład policja, sąd, protokołowanie),

© for this edition by CNS

2.5. uwzględnianie warstwy prozodycznej,
2.6. uwzględnianie warstwy niewerbalnej (gesty, mimika),
2.7. uwzględnianie interakcyjności (nakładanie się replik),
2.8. uwzględnianie segmentacji (na przykład głoski, sylaby, frazy intonacyjne).

3. Zastosowanie technologii:
   3.1. bezpośredni zapis mowy bez użycia nagrania;
   3.2. transkrypcja z nagrania:
      3.2.1. sprzęt amatorski,
      3.2.2. sprzęt profesjonalny,
      3.2.3. studio/laboratorium;
   3.3. transkrypcja na podstawie nagrania z wykorzystaniem programu do transkrypcji (ELAN, EXMARaLDA itp.);
   3.4. transkrypcja automatyczna przy użyciu programu do zamiany mowy na tekst (DictationPro, Voice Note itp.).
4. Ze względu na cel:
   4.1. badania fonetyczne,
   4.2. badania logopedyczne,
   4.3. analiza konwersacyjna,
   4.4. opis idiolektu, ustalanie autorstwa,
   4.5. dokumentacja mówienia (sąd, urząd, historia mówiona),
   4.6. artystyczny (literatura, publicystyka),
   4.7. mnemotechniczny (notatki terenowe w dialektologii, notatki z lekcji, z wykładów).

Podsumowanie

W artykule nieco bliżej zostały przedstawione typy transkrypcji, z którymi bardzo często spotykamy się nie tylko w naszej pracy naukowej, lecz także dydaktycznej. Młodzi badacze szczególnie na początkowym etapie pracy mają problemy w odróżnianiu zapisu tekstu pisanego od mówionego. Dominujące w świadomości językowej reguły związane z pisaną odmianą języka często blokują możliwość oddania cech wypowiedzi mówionej w ramach transkrypcji, co potwierdza jedna z metajęzykowych refleksji studenta na temat notowania samodzielnie sporządzonego nagrania: „Transkrypcja nie jest poprawna gramatycznie, zapisana jest tak, jak ją słyszymy, bez używania znaków interpunkcyjnych, bez przerw oraz poprawek". Istnieje więc jawna potrzeba pogłębionego opisu poszczególnych typów transkrypcji, aby szybciej i łatwiej zrozumieć ich istotę, sposób oraz cel zastosowania. Rozumienie różnic między poszczególnymi sposobami zapisu i tym samym analizy tekstu mówionego ułatwi oraz pomoże upowszechnić jego badania.

© for this edition by CNS

Bibliografia

Źródła

Czechowski R. (2020): *Pokoleniowe zmiany w języku komentatorów piłki nożnej*, nieopublikowana praca licencjacka, Wrocław.

Gaudy A. (2017): *Trudności artykulacyjne a błędy ortograficzne u osób rosyjskojęzycznych uczących się języka polskiego*, nieopublikowana praca magisterska, Wrocław.

Hordejuk G. (2020): *Mowa dzieci w sytuacji polsko-brytyjskiego bilingwizmu. Porównawcza analiza wybranych cech morfologicznych i leksykalnych*, nieopublikowana praca licencjacka, Wrocław.

www1: www.dialektologia.uw.edu.pl/index.php?l1=opis-dialektow&l2=dialekt-malopolski&l3=sadecczyzna&l4=sadecczyzna-gwara&l5=teksty-lososina-gorna1# dostęp: 13.07.2020).

Literatura

Czesak A., Okoniowa J., Waniakowa J. (2004): *Standard zapisu fonetycznego polskich tekstów gwarowych*, Kraków.

Dukiewicz L. (1995): *Fonetyka*, [w:] *Gramatyka współczesnego języka polskiego. Fonetyka i fonologia*, red. H. Wróbel, Kraków, s. 8–104.

Dunaj B. (red.) (1979): *Wybór tekstów języka mówionego mieszkańców Krakowa*, Kraków.

ISJP: Bańko M. (red.) (2000): *Inny słownik języka polskiego*, t. 1–2, Warszawa.

Jefferson G. (1984): *On Stepwise Transition from Talk about a Trouble to Inappropriately Next-positioned Matters*, [w:] *Structures of Social Action*, red. J.M. Atkinson, J. Heritage, Cambridge, s. 191–222.

Jefferson G. (2004): *Glossary of transcript symbols with an introduction*, [w:] *Conversation Analysis: Studies from the First Generation*, red. G.H. Lerner, Amsterdam, s. 13–31.

Levinson S.C. (2010): *Pragmatyka*, Warszawa.

Lorenc A. (2016): *Transkrypcja wymowy w normie i przypadkach jej zaburzeń. Próba ujednolicenia i obiektywizacji*, [w:] *Logopedia artystyczna. Logopedia XXI wieku*, red. B. Kamińska, S. Milewski, Gdańsk, s. 107–135.

Lubaś W. (red.) (1978): *Teksty języka mówionego mieszkańców miast Górnego Śląska i Zagłębia*, t. 1, Katowice.

Lubaś W (red.) (1980): *Teksty języka mówionego mieszkańców miast Górnego Śląska i Zagłębia*, t. 2, Katowice.

Majewska-Tworek A., Zaśko-Zielińska M., Pęzik P. (2020): *Polszczyzna mówiona miast — kontynuacja badań z lat 80. XX wieku z wykorzystaniem narzędzi lingwistyki cyfrowej*, „Forum Lingwistyczne", nr 7, https://www.journals.us.edu.pl/index.php/FL/article/view/10330.

Moore E., Llompart J. (2017): *Collecting, transcribing, analyzing and presenting plurilingual interactional data*, [w:] *Qualitative approaches to research in plurilingual language learning environments / Enfocaments qualitatius per a la recerca en educació plurilingüe*, red. A.E. Moore, M. Dooly, Dublin, s. 403–417.

Mróz P. (2016): *Zapis fonetyczny jako zjawisko typograficzne*, „Słowo. Studia Językoznawcze", nr 7, s. 196–207.

Porayski-Pomsta J., Kwasiborska-Dudek J., Jauer-Niworowska O., Lipiec D., Więcek-Poborczyk I., Golanowska M., Malinowska D. (2013): *Znaki fonetyczne do zapisu poprawnych i zdeformowanych realizacji fonemów języka polskiego w alfabetach międzynarodowym i slawistycznym*, Gdańsk.

PSPWP: Lubaś W., Urbańczyk S. (1994): *Podręczny słownik poprawnej wymowy polskiej*, Kraków-Katowice.

Rancew-Sikora D. (2007): *Analiza konwersacyjna jako metoda badania rozmów codziennych*, Warszawa.

© for this edition by CNS

Rogoziński Ł., Zarzeczny G. (2012): *Adapting GAT2 conventions for the transcription of spoken Polish as part of the GeWiss project*, „Studia Linguistica" 31, s. 69–78.
Rybka P. (2015): *Międzynarodowy alfabet fonetyczny w slawistyce*, Katowice.
Selting M., Auer P., Barden B., Bergmann J., Couper-Kuhlen E., Günthner S., Meier C., Quasthoff U., Schlobinski P., Uhmann S. (1998): *Gesprächsanalytisches Transkriptionssystem (GAT) / A Transcription System for Conversation Analysis*, „Linguistische Berichte" 173, s. 91–122.
SWFWP: Madelska L. (2005): *Słownik wariantywności fonetycznej współczesnej polszczyzny*, Kraków.
SWiONO: Bartmińska I., Bartmiński J. (1997): *Słownik wymowy i odmiany nazwisk obcych*, Bielsko-Biała.
SWP: Karaś M., Madejowa M. (red.) (1997): *Słownik wymowy polskiej*, Kraków.
WSJP: Żmigrodzki W. (red.): *Wielki słownik języka polskiego*, https: //www.wsjp.pl.
Zaręba A. (1955): *O metodach i technice badań gwarowych*, „Biuletyn Polskiego Towarzystwa Językoznawczego" 14, s. 140–155.
Zaśko-Zielińska M., Majewska-Tworek A., Śleziak M., Tworek A. (2020): *Od rozmowy do korpusu, czyli jak zbierać i archiwizować dane mówione*, Wrocław.

Functional description of transcription as a preliminary stage of spoken data analysis

Summary

The article presents main types of spoken text transcription, as well as describes the relationship between the research goal and transcription type selection. It also presents a transcription typology proposal. Most of the transcript examples are from the archives of the Oral Polish Laboratory.

Keywords: spoken data analysis, spoken text transcription, phonetics, dialectology, conversational analysis, lexicology, transcription typology.

© for this edition by CNS