

Z prac nad korpusem polsko- oraz rosyjskojęzycznych wiadomości przesyłanych na kartach pocztowych w drugiej połowie XX w. Zagadnienia transkrypcji i anotowania¹

DANIEL DZIENISIEWICZ, PIOTR WIERZCHOŃ
(Uniwersytet im. Adama Mickiewicza w Poznaniu)

1. WPROWADZENIE

Podstawą analiz językoznawczych jest tekst. Można go rozumieć różnorodnie: jako zapisy mowy, napisy miejskie, teksty literatury, publicystykę itd. W opisywanym przedsięwzięciu interesować nas będzie tekst kart pocztowych, które zostały włączone do obiegu (tzn. wysłane do adresata), ten artykuł zaś poświęcony jest aspektom cyfrowego przetwarzania tekstu pocztówki oraz procedurom operacyjnym zastosowanym przy tworzeniu przeszukiwalnej bazy kart pocztowych.

Zanim jednak przystąpimy do rozważań dotyczących konstruowania cyfrowej kolekcji interesujących nas obiektów oraz języka wiadomości wysyłanych za ich pośrednictwem, pragniemy określić, czym jest „karta pocztowa” oraz skrótkowo ukazać tło historyczne towarzyszące jej powstaniu. Zdzisław Wąsik w zwięzły sposób definiuje kartę pocztową jako kartę korespondencyjną zawierającą adres i stempel pocztowy, wysyłaną za pośrednictwem poczty. Stanisław

1 W językoznawstwie korpusowym pojęcie „korpus anotowany” oznacza zbiór tekstów uzupełniony o takie informacje jak np. dane o autorze, jego płci, pochodzeniu, wykształceniu, roku powstania, gatunku, a także dane metajęzykowe, np. etykiety części mowy, części zdania, cechy prozodyczne, dane semantyczne i leksykalne. Więcej na ten temat zob. w pracy Piotra Paryzka *Pozyskiwanie danych leksykalnych z tekstów elektronicznych (na materiale czasopisma naukowego)*, Poznań 2011, s. 33-34 (praca doktorska napisana pod kierunkiem Piotra Wierzchoń, dostępna pod adresem <https://repozytorium.amu.edu.pl/bitstream/10593/1422/1/Piotr%20Paryzek.pdf> – stan z 1 lutego 2018 r.).

Pietraszko stwierdza, że w potocznym rozumieniu pocztówka jest formą korespondencji, lecz z perspektywy teorii informacji stanowi środek przekazu informacji. Nie odróżnia jej to jednak od listu, który spełnia tę samą funkcję. Różnicę między listem a kartą pocztową stanowi natomiast publiczna dostępność tekstu pocztówki².

Pomimo istnienia wcześniejszych form korespondencji, takich jak bilety wizytowe, kartony świąteczne, kondolencyjne, zaproszenia na bale i polowania, druki ulotne i inne, które można by uznać za bezpośrednich przodków współczesnej pocztówki, pomysł stworzenia „otwartego listu” powszechnie kojarzony jest z obszarem niemieckojęzycznym i nazwiskiem Heinricha von Stephana, który jako pierwszy przedstawił ideę „pocztowego listka” (niem. *Postblatt*) 30 listopada 1865 r. w Karlsruhe na V Niemieckiej Konferencji Poczty. Jednak dopiero profesor ekonomii Akademii Wojskowej w Wiener-Neustadt Emanuel Hermann siłą swojego autorytetu sprawił, że 1 października 1869 r. Poczta Cesarstwa Austro-Węgierskiego wydała pierwsze karty pocztowe. W ciągu trzech miesięcy od wprowadzenia tego wynalazku w Austrii sprzedano ok. 3 milionów pocztówek. W 1870 r. austriacka inicjatywa zyskała popularność także w Niemczech, Szwajcarii, Luksemburgu i Wielkiej Brytanii. W następnych latach do grona państw korzystających z otwartego listu jako środka korespondencji dołączyły m.in. Belgia, Holandia, Dania, Finlandia, Kanada, Francja, Rosja, Szwecja, Norwegia, Cejlon, Chile, Stany Zjednoczone Ameryki, Hiszpania, Serbia, Rumunia, Włochy i Turcja³. Wkrótce pocztówka zyskała ogólnoświatową sławę, a szczyt jej popularności przypadł na lata 1900-1914, nazywane „złotym wiekiem pocztówki”⁴.

Dotychczasowe prace poświęcone kartom pocztowym w przeważającej mierze dotyczyły aspektów funkcjonowania tego medium niezwiązanych z językową stroną wiadomości wysyłanych za jego pośrednictwem, częstokroć skupiały się wyłącznie na warstwie ikonograficznej lub np. wydarzeniach historycznych, do których odsyłały informacje – zarówno graficzne, jak i tekstowe

2 Zob. S. Pietraszko, *Przekazy i wartości*, w: *Aksjosemiotyka karty pocztowej*, red. P. Banaś, Wrocław 1993, s. 39-40.

3 Informacje zawarte w tym akapicie pochodzą z pracy Jana Kotłowskiego, *Dawne pocztówki: historia, ikonografia, kolekcjonerstwo*, Warszawa 1998, s. 5-10. Zob. też pracę Ireny Zauchy, *Geneza i historia karty pocztowej. Problemy prawne i techniczne*, w: *Aksjosemiotyka...*, *op. cit.*, s. 63-75. Szybko upowszechniły się określenia pocztówki: niemieckie *Correspondenz-Karte*, *Postkarte*, francuskie *carte postale*, angielskie *post-card*, węgierskie *levelezési*, rosyjskie *открытое письмо (открытка)* i polska *karta pocztowa (pocztówka)*, a także synonimy: *korespondentka*, *widokówka* (dla kart ilustrowanych) – zob. *ibidem*, s. 5. Polskie jednowyrazowe określenie zostało wyłonione w drodze powszechnego konkursu, towarzyszącego wielkiej wystawie kart pocztowych w Warszawie w 1900 r. Jury, złożone z członków redakcji *Słownika polskiego*, wybrało spośród trzystu nadesłanych propozycji – pięć (*listówka*, *liścik*, *otwartka*, *pisanka* i *pocztówka*). Głosami zwiedzającej wystawę publiczności zwyciężyła właśnie *pocztówka* (zob. H. Duninówna, *Warszawskie nowinki 1815-1900*, Warszawa 1970, s. 365-366).

4 Zob. J. Szklarczyk-Mirecka, *Dzieje filokartystów w Polsce*, w: *Aksjosemiotyka...*, *op. cit.*, s. 79.

– zamieszczone na pocztówkach⁵. W szczególności należy jednak wskazać na brak wyczerpujących opracowań lingwistycznych prowadzonych na dużej liczbie tekstów i ukierunkowanych na ich kwantytatywną analizę. Paweł Banaś stwierdza:

- » kartą pocztową interesowano się głównie z racji jej wartości historyczno-dokumentacyjnych. Jako taka okazała się cennym źródłem dla etnografów, socjologów, historyków, a po części i historyków sztuki. Bywa także rozpatrywana jako samoistne dzieło sztuki oraz szczególny rodzaj *mass medium* – w tym ostatnim przypadku istotne są pytania nie tyle o ikonograficzną stronę przekazu, nie o to „co”, ale „jak” jest przekazane⁶.

Konkludując, badacz postuluje:

- » Pojedyncza karta pocztowa, dawny album, nowa kolekcja, ruch zbieracki i niezliczone motywy kart pocztowych mogą i powinny być przedmiotem badań różnych dyscyplin: historii i geografii, socjologii i estetyki, historii sztuki i historii kultury wreszcie⁷.

Pragnąc zatem uzupełnić oraz poszerzyć dotychczasową wiedzę na temat kart pocztowych, stawiamy sobie za cel sporządzenie wielopłaszczyznowego językoznawczego opisu wiadomości zgromadzonych w korpusie pocztówek polskich oraz radzieckich i rosyjskich nadanych w latach 1945–2000, uwzględniającego m.in. aspekty pragmatycznojęzykowe, socjolingwistyczne oraz historycznojęzykowe. Naszym zamiarem jest rejestracja i analityczny opis tekstu wiadomości w kontekście rozwoju Polski (ściślej: Polskiej Rzeczypospolitej Ludowej oraz III Rzeczypospolitej) i Rosji (a dokładniej: Związku Radzieckiego i Federacji Rosyjskiej) w latach 1945–2000.

5 O większym zainteresowaniu awersem, także w praktyce korespondencyjnej, świadczą następujące słowa: „Większa mobilność przyczyniła się do wzrostu zapotrzebowania na środki komunikacji i wymagała odpowiednich środków korespondencji. W porównaniu z listem na pocztówkach było o wiele mniej miejsca na korespondencję, można było pisać krótko albo wystarczyło zapełnić gotowe, już wydrukowane zdania. Dzięki temu wszystkie warstwy społeczne były w stanie korzystać z pocztówek. O wyborze pocztówki decydował awers, a więc motyw, natomiast sam tekst odgrywał rolę podrzędną i był często banalny [wyróżn. D.D., P.W.]” – zob. *Postkarten erzählen Geschichte: Die Stadt Posen 1896–1918 // Pocztówki opowiadają historię: Miasto Poznań 1896–1918*, red. S. Kemlein, Lüneburg 1997, s. 26.

6 P. Banaś, *Karta pocztowa jako przedmiot badań*, w: *Aksjosemiotyka...*, *op. cit.*, s. 11.

7 *Ibidem*, s. 13.

Celem artykułu jest przedstawienie wybranych zagadnień związanych z procesem tworzenia korpusu polsko- oraz rosyjskojęzycznych⁸ wiadomości wysyłanych za pośrednictwem kart pocztowych w drugiej połowie XX w. W tekście opisano etapy cyfrowego przetwarzania tekstu pocztówki, takie jak transkrypcja – rozumiana jako przepisanie tekstu z postaci papierowej na elektroniczną – oraz tworzenie etykiet (znaczników) przyporządkowywanych następnie poszczególnym elementom tekstu wiadomości, w skład których zaliczono m.in. zwrot do adresata, pozdrowienia, życzenia, formułę finalną i postscriptum. Wszystkie komponenty treści wiadomości poddano szczegółowej analizie. Omówiono przesłanki stojące za wyborem konkretnych – niekiedy kontrowersyjnych – procedur operacyjnych oraz zilustrowano je odpowiednimi przykładami. Prezentowane w artykule rozważania stanowią część projektu *100 000 minihistorii*, zakładającego budowę przeszukiwalnego pełnotekstowo korpusu ok. 100 000 polskich powojennych wiadomości przesyłanych na pocztówkach w latach 1945-2000⁹.

2. CHARAKTERYSTYKA POZYSKIWANEGO MATERIAŁU.

PROCES SKANOWANIA

Karty pocztowe nadane w latach 1945-2000 są pozyskiwane (kupowane) od osób prywatnych w liczbie od kilku egzemplarzy do kilku tysięcy sztuk, przede wszystkim za pośrednictwem Internetu, np. serwisów zamieszczających ogłoszenia o kupnie / sprzedaży oraz portali społecznościowych, na których funkcjonują grupy filokartystów (miłośników kart pocztowych), jak również w antykwariatach. Docelowo zgromadzona zostanie baza ok. 2500 polskojęzycznych oraz ok. 2500 rosyjskojęzycznych wiadomości. Należy przy tym zaznaczyć, że ze względu na szeroką dostępność oraz niską cenę pocztówek wyprodukowanych i nadanych po zakończeniu

8 W obecnym artykule opisy rosyjskojęzycznych wiadomości służą wyłącznie ukazaniu cech typowych dla tekstów pisanych w tym języku. Ze względu na fakt, że artykuł kierowany jest do odbiorców polskojęzycznych, przyjęto, iż podstawowy materiał cytawoty w tekście stanowią będą fragmenty wiadomości w języku polskim.

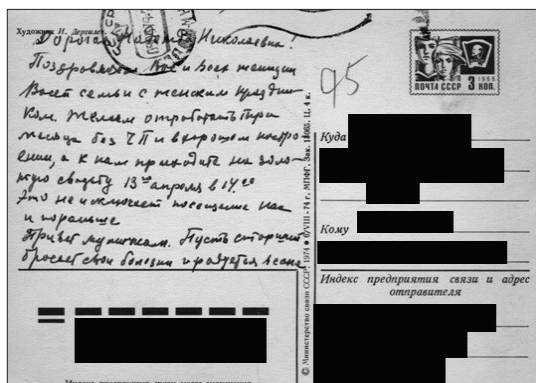
9 Te dane liczbowe i chronologiczne stanowią wstępne założenia projektu, w ramach którego każda pozyskana pocztówka zostanie zeskanowana oraz zapisana w formacie elektronicznym, a następnie co najmniej 50% tekstów wiadomości zostanie przepisanych i opatrzonych etykietami. Ponadto utworzone zostaną podkorpusy kart wakacyjnych oraz okolicznościowych. Zgromadzony materiał będzie największym tego typu zasobem przeznaczonym do badań nad językiem korespondencji prywatnej okresu PRL i III RP. Natomiast konstruowany korpus polsko- i rosyjskojęzycznych wiadomości to odrębny zasób, choć część wspólną obu przedsięwzięć stanowią wiadomości polskojęzyczne. Oprócz różnic językowych i proporcji zawartych w nich obiektów odmienne są także cele obu projektów: korpus polsko-rosyjski tworzony jest przede wszystkim z myślą o prowadzeniu analiz porównawczych, podczas gdy w ramach projektu *100 000 minihistorii* nadrzędnym celem jest stworzenie przeszukiwalnego zasobu *Big Data*.

drugiej wojny światowej tworzony korpus ma częściowo charakter oportunistyczny. Koszt przedwojennych pocztówek jest bowiem wysoki w porównaniu do kosztu pocztówek powojennych, co utrudnia zgromadzenie dużych partii materiału, niezbędnych do prowadzenia analiz ilościowych dla tego okresu. W kontekście badań nad korespondencją prywatną podobne trudności cechują również pozyskiwanie korespondencji listownej. Zaletą badań nad korespondencją pocztówkową jest fakt, że w wielu wypadkach możliwe jest wyodrębnienie z treści wiadomości takich informacji jak płeć nadawcy i adresata, ich liczba, łącząca ich relacja, lokalizacja czy data wysłania wiadomości, co znacząco ułatwia konstruowanie zasobu przeznaczonego do prowadzenia analiz socjolingwistycznych (zob. dalszą część artykułu). Zatem w przeciwieństwie do w pewnym stopniu oportunistycznych przesłanek stojących za wyborem okresu, z którego pochodzi materiał badawczy, dążymy do jego możliwie maksymalnego zrównoważenia w kontekście reprezentatywności ww. danych. W związku z tym podejmowane są próby zamieszczenia w korpusie porównywalnej liczby wiadomości pochodzących z poszczególnych dekad, wiadomości autorstwa zarówno kobiet, jak i mężczyzn, kart okolicznościowych, jak i wakacyjnych. Niewątpliwie w poczet mankamentów gromadzonego w ten sposób materiału należy zaliczyć fakt, iż w większości wypadków niemożliwe jest ustalenie takich metadanych jak wiek, pochodzenie społeczne bądź wykształcenie nadawców i adresatów. Ten stan rzeczy powodowany jest przez masowy, zapośredniczony sposób nabywania pocztówek, wykluczający bezpośredni kontakt z nadawcami i adresatami wiadomości, co uniemożliwia uzyskanie niektórych informacji dotyczących pozyskiwanych kart.

Po nabyciu kart pocztowych odbywa się pierwszy etap digitalizacji zgromadzonych materiałów: każda z tak pozyskanych pocztówek jest skanowana obustronnie w kolorze w rozdzielczości 300 dpi. Przykładowy efekt tej pracy został przedstawiony na poniższych fotografiach, ukazujących odpowiednio awers i rewers karty pocztowej:



Fot. 1. Awers karty pocztowej



Fot. 2. Rewers karty pocztowej

3. TRANSKRYPCJA I ANOTOWANIE

Następnie wiadomości są transkrybowane, a poszczególne elementy ich treści – anotowane. W tym celu stworzona została baza danych online, w której zamieszczane są zarówno fotografie awersów, jak i rewersów uprzednio zeskanowanych pocztówek. Ponadto utworzono interfejs, w którym po lewej stronie umieszczono skan awersu oraz rewersu pocztówki, po prawej zaś – pole do przepisywania treści wiadomości oraz panel etykiet służących do oznaczania poszczególnych elementów wiadomości.



Fot. 3. Interfejs do wykonywania transkrypcji i anotowania

System tworzenia i anotowania korpusu wykonany został z zastosowaniem popularnych języków skryptowych PHP i JavaScript oraz bazy danych MySQL. Komponent umożliwiający anotację wykonano na podstawie edytora WYSIWYG NicEdit, poprzez dodanie do obsługiwanych przez niego znaczników HTML ich

wariantów z dodatkowymi atrybutami, tak że w wyniku działania redaktora plik tekstowy otrzymuje postać dokumentu HTML, przykładowo:

```
<span class="greetings">  
  Moc pozdrowień  
  <span class="city">z Zakopanego</span>  
  ślą  
  <span class="signature">Irek i Darek</span>  
</span>  
<span class="wishes">  
  Oby Cię halny przywiał w nasze strony  
</span>
```

Na załączonym przykładzie zauważyć można m.in. pozdrowienia z zagnieżdżonymi wewnątrz zwrotami oznaczonymi jako nazwa miasta i podpis. Ze względu na fundamentalny charakter badań ukierunkowanych na wypracowanie zestawu kategorii treści karty pocztowej w celu możliwie kompleksowego skatalogowania specyficznej struktury tekstu wiadomości, dotychczas nieopisywanej z perspektywy informatycznej oraz językoznawczej, zdecydowano, że w przedsięwzięciu nie będą wykorzystywane funkcjonujące sposoby reprezentacji tekstu elektronicznego, jak np. standard *TEI* (*Text Encoding Initiative*). Nie wyklucza się jednak możliwości dostosowania ostatecznie wypracowanej struktury do istniejących standardów po zakończeniu prac.

Podczas przepisywania treści wiadomości do odpowiedniego pola w interfejsie w kolejności widniejącej na karcie pocztowej zachowywana jest ortografia oraz interpunkcja występująca w tekście oryginalnym, np. „Ja na tomiasz prawdo podobnie już w poniedziałek to jest 25.06.62 wypływam do Ameryki Północnej”. Nie uwzględnia się zaś takich parametrów oryginalnego tekstu jak zachowanie pierwotnej liczby wyrazów w każdym wierszu czy oznaczanie przenoszenia wyrazów. W przypadku wystąpienia problemów z odczytaniem fragmentów treści w miejsce nieczytelnych liter wstawiane są znaki zapytania, np. „Chyba będę siedzieć cały czas w pracy, bo pozostali ????? będą mieli urlopy”.

Treść stempla przepisywana jest wyłącznie wtedy, gdy informacje o dacie i lokalizacji wysłania pocztówki nie zostały zapisane odręcznie przez nadawcę. Ponadto nie zdecydowano się na utworzenie odrębnej etykiety do oznaczania stempla, gdyż zawiera on w swej treści informacje o dacie i lokalizacji wysłania karty, dla których przewidziano osobne znaczniki. Z kolei zaczerpnięte z treści wiadomości dane nadawcy i odbiorcy, takie jak imię, nazwisko i adres, są zamieszczane w specjalnie wyodrębnionych polach (zob. dalszą część artykułu).

W celu przeprowadzenia możliwie kompleksowej analizy wiadomości na podstawie wcześniejszego oglądu zgromadzonego materiału wyodrębniono powtarzające się w ich strukturze kategorie, które w dalszej kolejności posłużyły do utworzenia etykiet (znaczników), za pomocą których wyróżniane są odpowiednie fragmenty treści poprzez ich zaznaczenie i wybór właściwej opcji w interfejsie. Etykiety wypracowane na podstawie wstępnego oglądu tekstów w zdecydowanej większości przypadków umożliwiają skatalogowanie całej treści wiadomości, bez konieczności pomijania jakichkolwiek elementów. Utworzono następujący system znaczników:

1. *Zwrot do adresata,*
2. *Pozdrowienia,*
3. *Życzenia,*
4. *Święto / specjalna okazja,*
5. *Dodatkowa informacja,*
6. *Pytanie,*
7. *Podziękowanie,*
8. *Prośba,*
9. *Przeprosiny,*
10. *Formuła końcowa,*
11. *Podpis,*
12. *Postscriptum,*
13. *Lokalizacja,*
14. *Data,*
15. *Błąd,*
16. *Dopisek.*

3.1. Zwrot do adresata

Za pomocą etykiety *Zwrot do adresata* oznaczana jest formuła adresatywna występująca w pozycji inicjalnej w stosunku do treści całej wiadomości lub/i formuła adresatywna występująca w poszczególnych komponentach treści, np. „Droga Aniu!”, „Droga Pani Profesor!”, „Kochana Dorciu”, „Lucynko!”¹⁰. Analiza materiału wykazała, że elementami treści, w których formuły adresatywne pojawiają się najczęściej, są *Dodatkowa informacja*, *Przeprosiny*, *Podziękowanie* i *Prośba*:

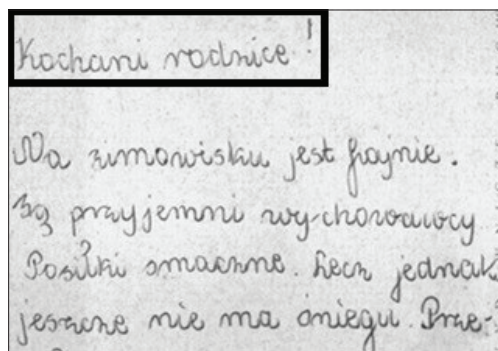
10 W przytaczanych fragmentach wiadomości zachowano ich oryginalną ortografię i interpunkcję. Szczegółowy opis wykładników lingwistycznych poszczególnych komponentów tekstu (aktów mowy), takich jak zwroty do adresata, pozdrowienia, życzenia itd., nie jest przedmiotem prezentowanego tekstu. Przytaczane są jedynie wybrane informacje ogólne.

1. „Kochana żoneczko, jeżdżę w Sanatorium ale jak długo będę to jeszcze nie wiem bo teraz pszechodzę badania” (*Dodatkowa informacja*);
2. „Kochanie. Wybacz mi że nie piszę listu, ale nie mam czasu” (*Przeprosiny*);
3. „Kochani, dziękujemy Wam za uroczyste przyjęcie oraz serdecznie zapraszamy w Bieszczady” (*Podziękowanie*);
4. „Kochana, odezwiąj się” (*Prośba*)¹¹.

Ogląd ogółu wiadomości pozwala również wysnuć wniosek, że zwrot do adresata może znajdować się:

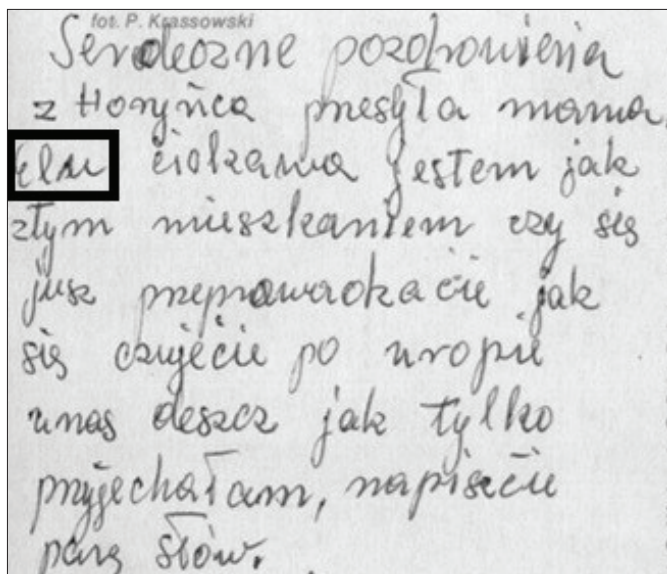
1. wyłącznie w pozycji inicjalnej w stosunku do całej wiadomości (zob. fot. 4.),
2. tylko jako część poszczególnych elementów treści (zob. wyżej *Dodatkowa informacja, Przeprosiny, Podziękowanie i Prośba* – zob. fot. 5.),
3. zarówno na początku wiadomości, jak i w treści jej poszczególnych elementów (zob. fot. 6.).

Sytuacja 3. ma miejsce wówczas, gdy wiadomość rozpoczyna formuła adresatywna skierowana do grupy osób, a w dalszej części następuje ograniczenie liczby odbiorców poprzez zawężenie treści komunikatu wyłącznie do jednej osoby, sygnalizowane przez użycie imienia odbiorcy w formie mianownika lub wołacza, np.: „Edmund”, „Elu”, „Zdziśku”. Sytuację 2. od sytuacji 3. odróżnia fakt, iż odbiorca nie jest w niej określony na początku wiadomości. Zwykle w tego typu tekstach zwrot do adresata w pozycji inicjalnej nie występuje w ogóle, a rozpoczyna je np. formuła pozdrowień bądź życzeń¹².

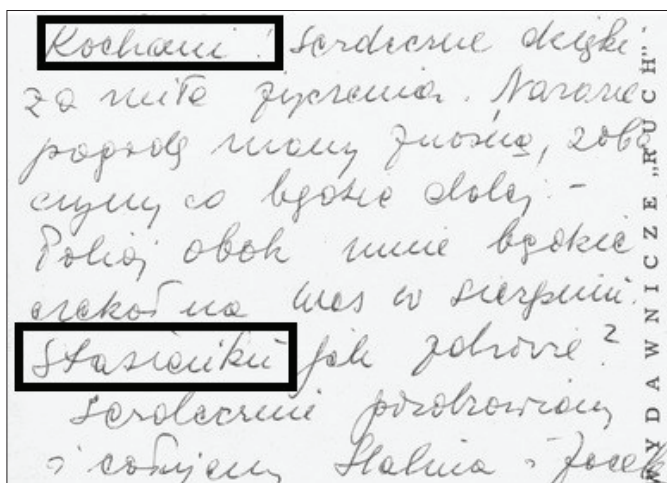


Fot. 4. Zwrot do adresata w pozycji inicjalnej w stosunku do całej wiadomości

- 11 W wypadku występowania kilku kategorii wewnątrz jednego, obszerniejszego elementu treści, etykiety nakładane są jedna na drugą, poczynając od elementów składających się z większej liczby wyrazów.
- 12 Więcej o aspektach poprawnościowych zwrotów do adresata i końcowych formuł grzecznościowych w listach oraz w wiadomościach przesyłanych za pośrednictwem pocztówek zob. w pracy Małgorzaty Marcjanik *Grzeczność w komunikacji językowej*, Warszawa 2013, s. 83-84.



Fot. 5. Zwrot do adresata poprzedzający pytanie



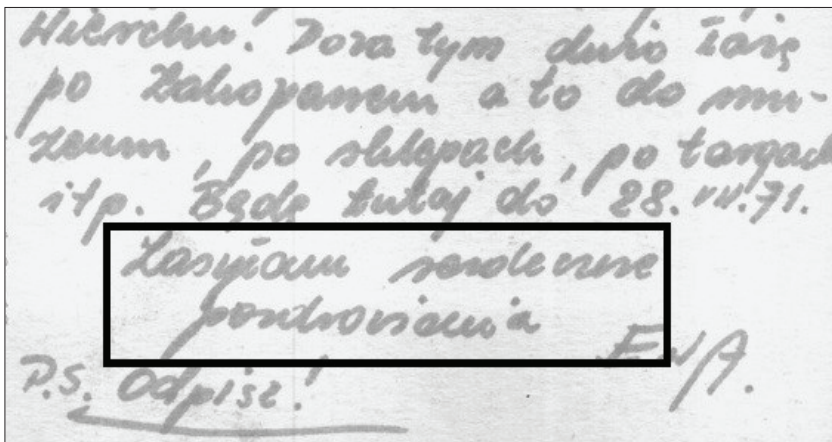
Fot. 6. Zwroty do adresata występujące na początku wiadomości oraz w treści pytania

3.2. Pozdrowienia

Znacznik *Pozdrowienia* odnosi się do formuł pozdrowień oraz wyrazów czułości, zawierających takie wyrazy, jak: „pozdrawiać”, „pozdrowienia”, „ściskać”, „uściski”, „całować”, „całusy”, np. „Pozdrowienia z Zakopanego”, „Serdecznie pozdrawiam Jadwiga”, „Sto całusów małym bobasom”, „uściski dla mamusi”. Wyróżniono pozdrowienia, w których:

1. nadawca i odbiorca nie są wymienieni (np. „Przesyłam pozdrowienia”),
2. wymieniony jest wyłącznie nadawca (np. „Serdeczne pozdrowienia z Kościeliska przesyłają Helenka z Goską”),
3. wymienieni są zarówno nadawca, jak i odbiorca (np. „Serdeczne pozdrowienia dla babci z Krakowa śle Ewa”).

Oprócz pozdrowień kierowanych przez nadawcę bezpośrednio do odbiorcy można wyróżnić pozdrowienia przekazywane osobom trzecim, znanym zarówno nadawcy, jak i odbiorcy, zob. np. sformułowanie „Pozdrowcie babcię i ciocię”. W wypadku gdy karty pocztowe przesyłane są z wyjazdu (np. wakacyjnego), w treści pozdrowień zazwyczaj umieszcza się nazwę miejscowości, z której kartka jest nadawana („Pozdrowienia z Zakopanego”). Pozdrowienia na ogół funkcjonują jako autonomiczna kategoria treści, lecz mogą też stanowić część innych kategorii (np. *Postscriptum*: „P.S. Osobne pozdrowienia dla Iwonki, Lucynki i Kasi”) bądź pełnić funkcję formuły końcowej poprzedzającej podpis (zob. niżej).



Fot. 7. Formuła pozdrowień w pozycji formuły końcowej

3.3. Życzenia

Etykieta *Życzenia* obejmuje formuły życzeń, które zazwyczaj funkcjonują jako samodzielna kategoria treści, np. „Przesyłam Ci moc gorących i serdecznych życzeń z Wrocławia”. Jednakże zdarzają się sytuacje, gdy życzenia łączą się z pozdrowieniami w obrębie jednej formuły, np. „W dniu Imienin zasyłam Najserdeczniejsze Życzenia oraz pozdrowienia z pięknych gór”. W treści życzeń częstokroć można wyodrębnić także komponent oznaczany etykietą *Święto / specjalna okazja*, np. formułę „Wesołych Świąt Dosiego Roku oraz wiele radości nad potomkiem życzy ciocia Freda”, w której występują określenia „Wesołych Świąt” (Święta Bożego

Narodzenia) i „Dosiego Roku” (Nowy Rok). Nie jest to jednak element obligatoryjny, gdyż można wskazać przykłady życzeń, w których nie występuje nazwa konkretnego święta, np. „Najserdeczniejsze (z możliwych) życzenia z Tatr w niezbyt typowej scenerii zimowej przesyła Jacek G”.

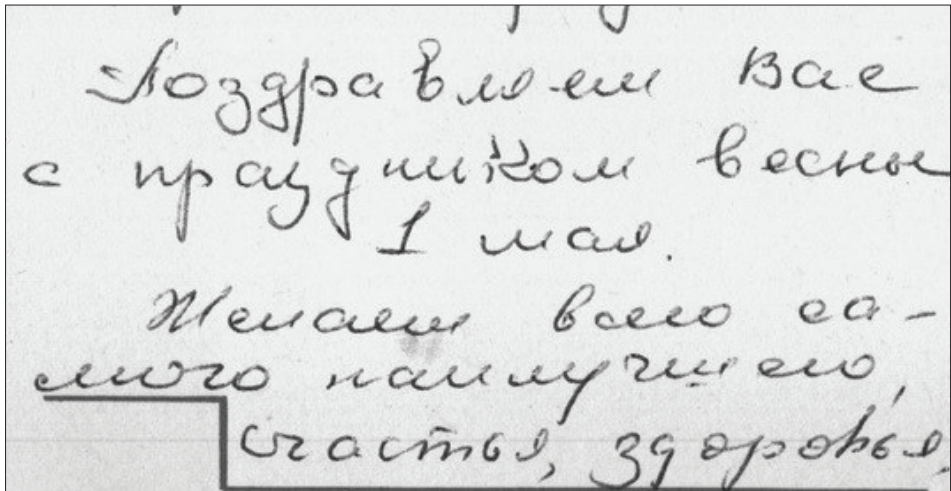
3.4. Święto / specjalna okazja

Jak wspomniano wyżej, ze znacznikiem *Życzenia* ściśle wiąże się etykieta *Święto / specjalna okazja*, odnosząca się do nazw świąt/okazji, z powodu których przesyłane są karty pocztowe, np. przykłady określeń zaczerpniętych z treści życzeń: „święt wielkanocnych”, „Nowego Roku”, „imienin”. Dotychczasowa analiza materiału pozwoliła na wyodrębnienie nazw świąt i okoliczności, które najczęściej występują w polsko- i rosyjskojęzycznych tekstach. Na tej podstawie kategoria *Święto / specjalna okazja* została podzielona na następujące podkategorie obejmujące poszczególne okoliczności:

1. *Dzień Kobiet*,
2. *Dzień Zwycięstwa*,
3. *Boże Narodzenie*,
4. *Święto Pracy*,
5. *Święto Rewolucji Październikowej*,
6. *Nowy Rok*,
7. *Urodziny*,
8. *Wielkanoc*,
9. *Imieniny*,
10. *Inne*¹³.

W sprawie budowy formuł życzeń warto nadmienić, że o ile w tekstach polskojęzycznych zwykle przyjmują one formę jednozdaniową, to w tekstach rosyjskich zazwyczaj obejmują dwa zdania. Pierwsze zdanie składa się z konstrukcji (lub jej wariantu) „*поздравлять с*” (pol. „wszystkiego najlepszego z okazji”) + nazwa święta, drugie zaś rozpoczyna się od czasownika „*желать*” (pol. „życzyć”), po którym wymieniane są konkretne wartości stanowiące treść życzeń, np. „*Поздравляю Вас с Новым 1978 годом. Желаю вам хорошего здоровья, отличного настроения, а самое главное долгих лет жизни*”.

13 Podkategoria *Inne* obejmuje wszelkie okoliczności występujące w materiale badawczym, których nie można sklasyfikować na podstawie pozostałych podkategorii, np. Dzień Nauczyciela – np. formuła „*с Днём Учителя!*” (tłum. „Wszystkiego najlepszego z okazji Dnia Nauczyciela!”).



Fot. 8. Formuła życzeń w rosyjskojęzycznej wiadomości

Zagnieżdżenie etykiety *Święto / specjalna okazja* w obrębie etykiety *Życzenia* umożliwia automatyczne wyodrębnianie formuł życzeń składanych z powodu konkretnych okoliczności. Zabieg ten pozwala na szybkie i efektywne zbadanie ich cech dystynktywnych (m.in. inwariantów formuł czy wartości występujących w treści życzeń składanych z konkretnej okazji). Zdajemy sobie sprawę, że współwystępowanie w korpusie takich etykiet jak *Święto / specjalna okazja* oraz np. *Pozdrowienia* może powodować wrażenie ich nierównoważności. Obecność omawianej etykiety, a w szczególności wyodrębnione w jej ramach podkategorie, jak np. *Boże Narodzenie* czy *Dzień Kobiet*, może kontrastować z takimi znacznikami jak np. *Pytanie*, *Prośba* czy *Dodatkowa informacja*, w których to nie wyodrębnia się podkategorii tematycznych, jak np. *Pytanie o zdrowie*, *Prośba o pożyczkę* czy *Informacja o warunkach w pensjonacie*. Taki zabieg byłby bowiem niepraktyczny ze względu na dużą różnorodność tematyczną treści tych kategorii. Dlatego zdecydowano, że dalsze analizy leksykalno-tematyczne w obrębie tych elementów będą na początkowym etapie badań prowadzone bez wykorzystania metod automatyzujących pracę z korpusem (anotowania tematycznego). Postanowiono jednak zachować kategorię *Święto / specjalna okazja* wraz z podkategoriami ze względu na kulturowo ograniczony, skończony i przewidywalny katalog świąt.

3.5. Dodatkowa informacja

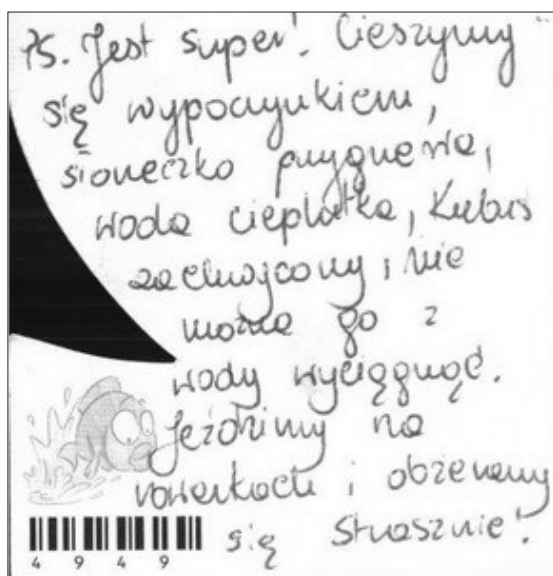
Znacznik *Dodatkowa informacja* obejmuje wszelkie zdania w treści wiadomości, w których nadawca przekazuje adresatowi informację (zdania oznajmujące). Zaliczono do nich zdania, które nie zawierają elementów charakterystycznych dla

pozdrowień, życzeń, prośb, podziękowań, pytań bądź formuł końcowych. Ta negatywna definicja obejmująca swym zasięgiem wiele zróżnicowanych aktów mowy, mimo swych mankamentów, jest jednak użyteczna na obecnym etapie konstruowania korpusu, ponieważ tematyczna klasyfikacja treści w ramach etykiety *Dodatkowa informacja* odbywać się będzie na podstawie manualnego oglądu wchodzących w jej skład zdań, bez wykorzystania metod automatyzujących. Za pomocą omawianej etykiety oznaczono np. następujące fragmenty:

1. „Na zimowisku jest fajnie. Są przyjemni wychowawcy Posiłki smaczne. Lecz jednak jeszcze nie ma śniegu”;
2. „Przyjechałam szczęśliwie, ulewny deszcz cały dzień. Teraz pogoda cudowna. Przykrzy mi się bez was wszystkich”;
3. „Deszcz pada, rzadko świeci słońce. W programie 10 wycieczek. Już jestem po trzech. Morskie Oko, spływ Dunajcem przede mną. Kurtka bardzo się przydaje. „Fajowo” jest na czasach”.

Często fragmenty oznaczone jako *Dodatkowa informacja* stanowią treść *postscriptum* (jest to charakterystyczne także dla pytań, prośb i podziękowań – zob. dalej), np.:

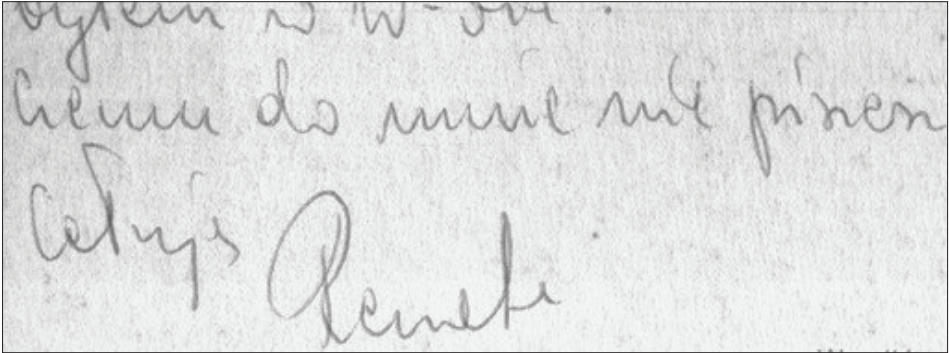
1. „PS Kurs kończy się 8. VI. Ja wracam 5 wieczorem i 7. VI będę na zajęciach”;
2. „PS. Jest dobrze. Wczoraj lało a dziś pogoda dla bogaczy ok jest wiatr i trochę kołysz. Grupa zdyscyplinowana, program b. wypełniony, brak czasu wolnego Wyjeżdżamy dzisiaj do Stargardu Gd. na noc a jutro do Malborka”;
3. „P.S. Pogoda fantastyczna. Jestem jak mulatka. Wieczory raczej udane”.



Fot. 9. Fragment oznaczony jako *Dodatkowa informacja* w treści *postscriptum*

3.6. Pytanie

Etykieta *Pytanie* oznaczane są pytania stawiane adresatowi przez nadawcę. Od strony graficznej znacznik ten obejmuje wszystkie zdania zakończone znakiem zapytania i/lub mające szyk zdania pytającego – zaobserwowano bowiem wypadki niezapisywania znaku zapytania na końcu wiadomości, np. „Czemu do mnie nie piszesz”. Oto przykłady pytań: „Jestem ciekawa czy otrzymaliście pożyczkę na budowę?”, „Czy Tatuś wziął zastrzyki?”, „Dlaczego nic nie piszesz?”¹⁴.



Fot. 10. Pytanie niezakończone znakiem zapytania

3.7. Podziękowanie

Znacznik *Podziękowanie* obejmuje formuły podziękowań, np. „Dziękujemy za pocztówkę, którą dzisiaj otrzymaliśmy”, „Dziękuję bardzo Cioci za pomoc w pilnowaniu Babci”, „Dziękuję za babkę była świetna”¹⁵.

3.8. Prośba

Etykieta *Prośba* stosowana jest do oznaczania prośb, np. „Napisz co u Ciebie”, „Przyjedźcie do mnie w niedzielę”, „przywiozę od p. Reni lekarstwo – proszę odebrać i przynieść grzyby”¹⁶.

14 Zob. także pytania występujące w treści *postscriptum*: „P.S. Jadziu dlaczego nic nie piszesz?”, „PS. Jak Ci się wiedzie po koloniach?”.

15 Zob. także podziękowania występujące w treści *postscriptum*: „PS. Bardzo dziękuję za życzenia sprawiły mi wielką radość że ktoś jeszcze pamięta”, „P.S. Dziękujemy za widokówkę”, „PS. Dziękuję bardzo za życzenia imieninowe”.

16 Zob. również prośby występujące w treści *postscriptum*: „P.S. Proszę przesłać pozdrowienia dla Szo-tów i Nowaków”, „P.S. Bardzo proszę o sprawdzenie, czy zostało mi wysłane zaświadczenie pracy i opinia, bo muszę to przedłożyć w Dziekanacie”, „P.S. Proszę podać swój nowy adres, to przyślę Panu nowe znaczki”.

3.9. Przeprosiny

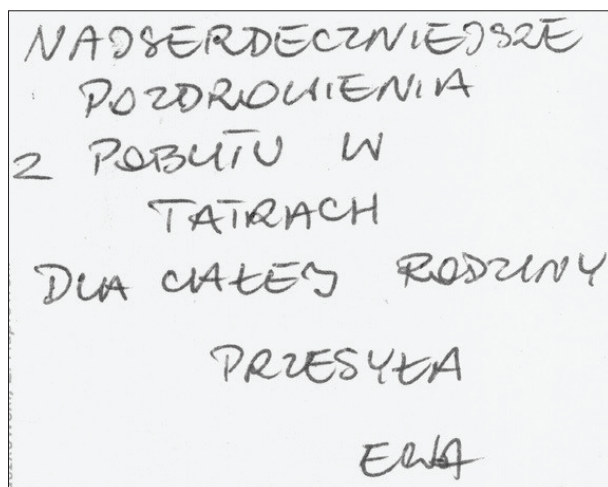
Kategoria *Przeprosiny* obejmuje formuły przeprosin zawierające czasowniki „przepraszać” lub „wybaczać”, np. „Bardzo przepraszam że nie zawiadamiam Pani osobiście, lecz jestem już na wakacjach”, „Aniu wybacz mi, że nie mogłem przyjść w sobotę”, „Przepraszam za pismo, ale piszę na kolanie”¹⁷.

3.10. Formuła końcowa

Znacznik *Formuła końcowa* odnosi się do formuły zamykającej wiadomość, znajdującej się bezpośrednio przed podpisem, np. „Do zobaczenia”, „No to na razie!”, „Hej!”. Jako formuły końcowe często wykorzystywane są formuły pozdrowienia, np. „Całuję mocno”, „Całuję Was mocno i pozdrowienia jeszcze raz”, „Jeszcze raz gorąco pozdrawiam i ściskam wszystkich”. Również w wiadomościach rosyjskojęzycznych pozdrowienia stanowią częsty sposób kończenia wiadomości, zob. „С приветами”, „С огромным приветом”, „Передайте привет и поздравления тёте Руте, дяде Володе, Лиде и Коле”.

3.11. Podpis

Etykieta *Podpis* służy do oznaczania podpisu pod wiadomością, np. „Anka”, „Irek i Darek”, „Asia z rodzicami”. Podpis może stanowić odrębną część wiadomości (zob.



Fot. 11. Podpis stanowiący część formuły pozdrowień

17 Zob. także przeprosiny występujące w treści *postscriptum*: „Ps Przepraszam że tak długo nic nie napisałem, to z przyczyny chodzenia na uszach. Pozdrówcie p (S)”, „P.S. Przepraszam za piątek nie mogłam przyjść, ponieważ wyjechałam w czwartek do Wałbrzycha”, „P.S. Bardzo przepraszamy że znowu zapomnieliśmy o dniu 26.VII”.

powyższe przykłady) lub znajdować się w pozycji podmiotu w zdaniu zawierającym pozdrowienia, np. „Najserdeczniejsze Pozdrowienia z pobytu w Tatrach dla całej rodziny przesyła Ewa”. Za potraktowaniem imienia „Ewa” jako podpis w powyższej formule przemawia to, że zostało ono graficznie wyodrębnione z tekstu pozdrowień przez przeniesienie do nowej linii.

3.12. Postscriptum

Znacznik *Postscriptum* służy do oznaczania *postscriptum*, tj. dopisku do listu, umieszczonego po podpisie autora. W skład *postscriptum* mogą wchodzić dodatkowe informacje, podziękowania, prośby, pytania, przeprosiny oraz pozdrowienia bądź prośby o pozdrowienie osób trzecich, np.:

1. „P.S. Towarzyszy nam deszcz, słońce nawet grad, ale w Schronisku u Krzeptowskich smak szarlotki uwieńczył nasze zmagania z aurą”;
2. „P.S. Pogoda w kratkę. Miejscowość bardzo ładna. Mimo tego bardzo się nudzę ponieważ na moim turnusie są same pryki. Wracam 29.06. Może się jeszcze spotkamy przed moim wyjazdem. Serdeczne pozdrowienia dla mamusi, tatusia i Marka”;
3. „P.S. Pozdrowienia dla cioci, wujka, Grzegorza Witalisa i Danki”;
4. „P.S. Proszę pozdrowić wszystkich”.

3.13. Lokalizacja

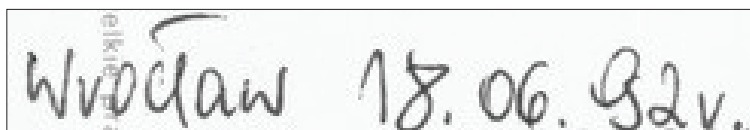
Etykieta *Lokalizacja* obejmuje nazwę miejscowości (miejsca), z której wysyłana jest karta pocztowa. Nazwa miejscowości może:

1. być zapisana odręcznie, zazwyczaj w pozycji poprzedzającej datę, np. „Zakopane 13.09.74”;
2. występować bezpośrednio w tekście wiadomości: „Serdeczne pozdrowienia dla babci z Krakowa śle Ewa” (za pomocą etykiety *Lokalizacja* został oznaczony tylko fragment „z Krakowa”);
3. znajdować się na stemplu pocztowym.

Nierzadkie jest także występowanie nazwy miejscowości we wszystkich zaprezentowanych powyżej pozycjach na jednej karcie pocztowej. Przyjęto, że w takiej sytuacji za pomocą etykiety *Lokalizacja* oznaczana jest nazwa zapisana odręcznie. Jeśli zapis odręczny nie występuje, nazwa miejscowości przepisywana jest bezpośrednio ze stempla wielkimi literami w nowym wierszu, poza głównym tekstem wiadomości, a następnie oznaczana za pomocą etykiety *Lokalizacja*, np. „CIECHOCINEK”, „ZAKOPANE”.

3.14. Data

Znacznik *Data* odnosi się do daty zapisanej w treści wiadomości, np. „Szczecin 11.7.1988 r.”, bądź znajdującej się na stemplu pocztowym. Podobnie jak w wypadku etykiety *Lokalizacja*, w pierwszej kolejności oznaczana jest odręcznie napisana data występująca w treści wiadomości, a przy jej braku – data ze stempla. W ostatnim przypadku cyfry przepisywane są w nowym wierszu, poza głównym tekstem wiadomości, ściśle według porządku, w którym widnieją na stemplu, np. „2567916” (25 czerwca 1979 r.).



Fot. 12. Miejscowość i data dopisane odręcznie

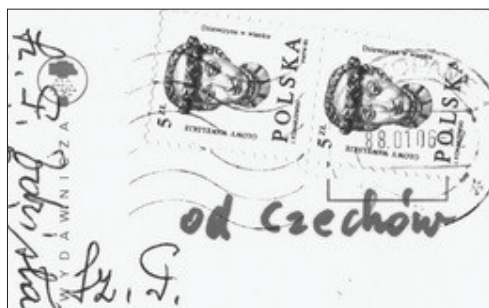
3.15. Błąd

Etykieta *Błąd* służy do oznaczania omyłek ortograficznych występujących w treści wiadomości. Dotychczas wyodrębniono następujące ich typy:

1. błędy w pisowni łącznej, rozdzielnej i w użyciu łącznika (np. „Codzień”, „dowidzenia”, „po obiednich”, „z-tąd”),
2. błędy w pisowni wielkich i małych liter (np. „gubałówce”, „ustce”),
3. błędy polegające na zastępowaniu, pomijaniu i dopisywaniu określonych liter (np. „Monachjum”, „nieprętko”, „Moji”).

3.16. Dopisek

Znacznik *Dopisek* dotyczy tekstu (w założeniu) dopisanego do wiadomości *post factum*, tj. już po przesłaniu karty pocztowej. Notatki tego typu wykonywane są najczęściej innym kolorem długopisu, flamastrem, kredką bądź ołówkiem przez osobę inną niż autor wiadomości.



Fot. 13. Tekst dopisany flamastrem

Ponadto utworzono pola *Nadawca* i *Adresat* służące do przepisywania danych nadawcy i adresata ze specjalnych rubryk wyodrębnionych na karcie pocztowej lub bezpośrednio z treści wiadomości (imiona, niekiedy nazwiska) w wypadku niewpisania tych informacji do ww. rubryk przez nadawcę.

Nadawca	Adresat
##Ewa##	Sz. P. [Redacted]

Fot. 14. Pola *Nadawca* i *Adresat*

Utworzono także pola *Płeć nadawcy* oraz *Płeć adresata*. Informacje te ustalane są na podstawie imion widniejących w treści wiadomości bądź – w wypadku ich braku – form fleksyjnych wyrazów. Opracowano następujący umowny system znaczników:

1. *k* – kobieta;
2. *m* – mężczyzna;
3. *lmk* – dwie lub więcej kobiet;
4. *lmm* – dwaj lub więcej mężczyzn;
5. *lmmk* – grupa mieszana.

Wprowadzone zostały również pola *Liczba nadawców* oraz *Liczba adresatów*, które to informacje pozyskiwane są z rubryk przeznaczonych do wpisywania danych nadawcy i adresata bądź – w wypadku ich braku – ustalane są na podstawie liczby podpisów pod wiadomością lub na podstawie liczby osób pozdrawiających/życzących oraz liczby osób, do których kierowane są pozdrowienia/życzenia.

Pole *Znormalizowana data* służy do wpisywania daty zaczerpniętej z treści wiadomości bądź ze stempla pocztowego, która jest następnie sprowadzana do postaci czterocyfrowej, np.:

1. 88.01 06 → 1988;
2. 21.12.72 → 1972;
3. 15 VII 78 → 1978.

Natomiast pola *Znormalizowana nazwa miejscowości nadawcy* i *Znormalizowana nazwa miejscowości odbiorcy* służą do wpisywania nazw miejscowości sprowadzonych do formy kanonicznej, np.:

1. „z Krakowa” → Kraków;
2. „Zakopanego” → Zakopane;
3. „w Ciechocinku” → Ciechocinek.

Poza tym wprowadzono kategorię *Relacja*, która określa relację łączącą nadawcę i adresata, ustalaną na podstawie lektury treści wiadomości. Opracowano następujący umowny system znaczników:

1. *r* – relacja rodzinna;
2. *k* – relacja koleżeńska;
3. *p* – relacja pracownicza;
4. *i* – relacja innego typu (np. rozwiązania krzyżówek wysyłane do redakcji gazet na kartach pocztowych).

Powyższe operacje normalizacyjne są niezbędne dla późniejszej automatyzacji ekscerpcji danych oraz do tworzenia statystyk.

Relacja: <input type="text" value="r"/>	Znormalizowana data: <input type="text" value="1989"/>
Płeć nadawcy: <input type="text" value="k"/>	Znormalizowana nazwa miejscowości nadawcy: <input type="text" value="Zakopane"/>
Płeć adresata: <input type="text" value="m"/>	Znormalizowana nazwa miejscowości odbiorcy: <input type="text" value="Warszawa"/>
Liczba nadawców: <input type="text" value="1"/>	
Liczba adresatów: <input type="text" value="1"/>	

Fot. 15. Pola *Relacja*, *Płeć nadawcy*, *Płeć adresata*, *Liczba nadawców*, *Liczba adresatów*, *Znormalizowana data*, *Znormalizowana nazwa miejscowości nadawcy*, *Znormalizowana nazwa miejscowości odbiorcy*

4. PERSPEKTYWY BADAWCZE

Tak skonstruowany zestaw kategorii pozwala na szybkie uzyskanie wyników, np. statystyk występowania formuł adresatywnych, pozdrowień, życzeń, formuł końcowych, dodatkowych informacji, pytań, prośb, przeprosin czy podziękowań. Dzięki mechanizmom wyszukiwania możliwe jest ustalenie częstości użycia poszczególnych formuł.

Ponadto system umożliwia poszukiwanie wybranych kategorii z poziomu innych kategorii za pomocą pola *Filtruj*. Jego działanie najlepiej zaprezentować na przykładzie. Załóżmy, że użytkownik bazy chce prześledzić, jakie formuły adresatywne występują w wiadomościach pisanych przez kobiety do mężczyzn. W tym celu należy ograniczyć wyszukiwanie na poziomie etykiety *Zwrot do adresata* do wiadomości, które w polu *Płeć nadawcy* zostały oznaczone za pomocą litery *k*, a w polu *Płeć adresata* za pomocą litery *m*. W analogiczny sposób można wyszukiwać inne konfiguracje płci nadawców i adresatów.

Wykorzystanie opisanego wyżej narzędzia umożliwia np. wydzielenie formuł życzeń związanych z konkretną okolicznością, np. życzeń imieninowych bądź noworocznych (zob. znacznik *Święto / specjalna okazja*), i zbadanie ich pod kątem stałości składu leksykalnego. Ponadto np. eksploracja korpusu z poziomu etykiety *Postscriptum* pozwala na oszacowanie, jakie elementy treści występują najczęściej w tej kategorii (np. *Dodatkowa informacja*, *Pytanie*, *Prośba*, *Podziękowanie*). Taki mechanizm wyszukiwania w korpusie pozwala również np. na sprawdzenie, z jaką częstością w miejsce formuły końcowej stosowane są formuły pozdrowień („całuję”, „uściski”, „pozdrowienia” itp.). Natomiast filtrowanie informacji chronologicznych umożliwia prześledzenie zmian, które zachodziły na przestrzeni lat, np. w użyciu określonych formuł adresatywnych bądź formuł pozdrowień (np. sformułowania „zasiłam pozdrowienia”, „przesyłam pozdrowienia”, „pozdrawiam”, „pozdrowienia”). Znormalizowane wiadomości geograficzne pomagają zaś ustalić, czy niektóre cechy występujące w wiadomościach mogą mieć charakter dialektalny bądź środowiskowy (np. zapis „dołancza”). W pozostałych wypadkach zawężenie wyszukiwania może stanowić pierwszy etap kolejnych analiz, np. klasyfikacji tematycznej fragmentów oznaczonych za pomocą znaczników *Dodatkowa informacja* bądź *Pytanie*, występujących wewnątrz kategorii *Postscriptum*.

W przyszłości korpus mógłby zostać uzupełniony o informacje o rodzajach kart pocztowych (np. karty świąteczne i okolicznościowe, karty wysyłane podczas wyjazdu wakacyjnego, wycieczki itp.) oraz dotyczące korelacji między treścią wiadomości i znajdującym się na awersie elementem ikonograficznym, które również podlegałyby możliwości wyszukiwania.

Podsumowując powyższy opis perspektyw badawczych, należy stwierdzić, że korpus polsko- oraz rosyjskojęzycznych wiadomości przesyłanych na kartach pocztowych zaopatrzone w omówione narzędzia może stanowić unikalny wstęp (bazę empiryczną) do badań o charakterze socjolingwistycznym, przede wszystkim zaś badań z zakresu lingwistyki płci oraz dialektologii. Ponadto planowane analizy obejmą zagadnienia etykiety językowej, tj. użycia formuł grzecznościowych w tekstach wiadomości oraz ich ewolucję (znacznik *Data*).

Należy nadmienić, że korpus nie został jeszcze opublikowany, jednak planowane jest jego udostępnienie po zakończeniu prac związanych z gromadzeniem, transkrypcją oraz anotowaniem tekstów. Obecnie dostęp do wersji roboczej zasobu mają wyłącznie współpracownicy projektu.

5. PODSUMOWANIE

Dotychczasowe zainteresowanie kartą pocztową koncentrowało się głównie na analizie jej strony graficznej bądź na jej tekście silnie skorelowanym z konkretnym miejscem, wydarzeniem, zagadnieniem. Nasze potrzeby poznawcze są odmienne: interesuje nas analiza pocztówek w ich maksymalnej do pozyskania i badawczego objęcia masie tekstowej. Tworzony korpus pozwoli na wykrycie serii specyficznych własności, jakie wylaniają się z analizy dużej liczby tekstów pocztówek. Wyodrębnione kategorie tekstu pocztówki (będące podstawą do sporządzenia syntetycznego opisu jego struktury) mogą stanowić uniwersalną bazę dla analogicznych badań prowadzonych na materiale innych języków, która mogłaby być dowolnie modyfikowana w zależności od potrzeb badawczych oraz konwencji występujących w praktyce piśmienniczej użytkowników konkretnego języka.

Badania frekwencji wyrazów, kolokacji i utartych formuł fatycznych pozwolą, z jednej strony, ustalić stopień schematyczności w myśleniu autorów wiadomości, z drugiej – stopień ich zindywidualizowania, umiejętności wyjścia poza sztywne ramy narzuconych przez uzus form językowych. Możliwe będzie także dokładne zbadanie na reprezentatywnym materiale empirycznym, w jakim stopniu pocztówki jedynie podtrzymują kontakt między osobami komunikującymi się, a na ile przekazują realne informacje, dostarczając odbiorcy wiedzę o otaczającej go rzeczywistości.

Nie bez znaczenia będą badania nad tematyką wiadomości. Ten aspekt wydaje się szczególnie interesujący w kontekście przemian politycznych i społecznych, jakie zaszły w Polsce (PRL i III RP) oraz Rosji (Związek Radziecki i Federacja Rosyjska) w latach 1945-2000. Analiza tekstów wysyłanych za pośrednictwem kart pocztowych pozwoli prześledzić zmiany zachodzące w obu językach (lub ich brak) w następstwie transformacji ustrojowych, tj. zmierzchu Polski Rzeczypospolitej Ludowej oraz Związku Socjalistycznych Republik Radzieckich i powstaniu odpowiednio III Rzeczypospolitej oraz Federacji Rosyjskiej.

Warto nadmienić, że dotychczas interesujące rezultaty przyniosła wstępna analiza porównawcza stopnia odtwarzalności poszczególnych komponentów treści wiadomości. W wiadomościach pisanych w języku polskim zauważalna jest

bowiem większa różnorodność stosowanych środków językowych (po przeprowadzeniu oglądu elementów treści podlegających anotowaniu zauważono także częstsze przypadki występowania komizmu językowego) w badanych formułach, a co za tym idzie – mniejsza schematyczność treści, zob. np. poniższe formuły życzeń:

„Serdeczne życzenia Wielkanocno-jajeczno-wiosenne ślą [podpis]”.

„Поздравляю я вас с праздником 1м мая и 9м мая. Желаю всем хорошего здоровья успеха в труде и самого хорошего праздничного настроения и пусть всегда будет мир”¹⁸.

Ponadto należy zwrócić uwagę na różnice występujące w treści przekazywanych informacji (etykieta *Dodatkowa informacja*). W tekstach pisanych po polsku odnaleziono dotąd bowiem więcej elementów opatrzonych tym znaczkiem, są one również dłuższe i zawierają więcej informacji, a także – tak jak w przypadku przywoływanej formuły życzeń – częściej zawierają elementy komizmu językowego oraz charakteryzują się pozytywnym wydźwiękiem (wniosek ten ma charakter globalny i odnosi się do całości materiału, bez względu na zastosowane filtry), np.:

» Właśnie odpoczywam w parku po podróży. Park pełen jest pięknych ptaków co zauważyłem gdy narobiły na mnie z lotu bądź też z pięknych drzew rosnących wokół. Podróż miałem na ogół dobrą, szkoda tylko, że złamali mi 2 żebra, lewą rękę i odbili prawe płuco. Jestem trochę głodny, bo mija 3 dzień jak odpoczywam sobie tutaj po podróży, a zostały mi tylko landrynki, bo kanapki ukradli mi razem z plecakiem. Ale to było jeszcze w Łodzi. Jeszcze trochę odpocznę i za chwilę się tutaj rozejrzę. Może nawet wślizgnę się do Panoramy Raławickiej.

Natomiast ich rosyjskie odpowiedniki są zwykle krótsze i częściej cechuje je wymowa negatywna, np.:

„С трудом пишу – болит рука и совсем отказывает мне служить. В 76ом году нам всем не повезло – все бодем”¹⁹.

Powyższe wnioski – sformułowane na podstawie wstępnego rozeznania – po przeprowadzeniu dokładniejszych badań mogą stanowić podstawę do dalszych analiz nad podobieństwami i różnicami kulturowymi wybranych obszarów języ-

18 Tłum. „Wszystkiego najlepszego z okazji świąt 1 i 9 Maja. Życzę wszystkim dobrego zdrowia, sukcesów w pracy, najlepszego świątecznego nastroju i niech zawsze będzie pokój”.

19 Tłum. „Z trudem piszę – boli mnie ręka i zupełnie odmawia mi posłuszeństwa. W siedemdziesiątym szóstym nie mieliśmy szczęścia – wszyscy chorujemy”.

kowych odzwierciedlonymi w tekstach prywatnej korespondencji. Należy bowiem zaznaczyć, że wyniki badań prowadzonych nad tekstami napisanymi w języku polskim i rosyjskim na tym etapie rozwoju obu państw (PRL i ZSRR) mogą ukazać istotne podobieństwa i różnice w prywatnym komunikowaniu się przedstawicieli dwóch kultur żyjących w systemach politycznych o zbliżonym charakterze (m.in. – w różnym stopniu – stosujących cenzurę, ograniczających wolność wypowiedzi i poglądów czy wolność religijną) lecz funkcjonujących na obszarach, w których historycznie przeważały pierwiastki właściwe dla różnych kręgów cywilizacyjnych – łacińskiego i bizantyńsko-turańskiego²⁰.

BIBLIOGRAFIA:

Aksjosemiotyka karty pocztowej, red. P. Banaś, Wrocław 1993, tu zwłaszcza:

Banaś P., *Karta pocztowa jako przedmiot badań*, s. 5-15;

Pietraszko S., *Przekazy i wartości*, s. 29-47;

Szklarczuk-Mirecka J., *Dzieje filokartystów w Polsce*, s. 77-87;

Wąsik Z., *Karty pocztowe w perspektywie semiologicznej*, s.17-28;

Zauchna I., *Geneza i historia karty pocztowej. Problemy prawne i techniczne*, s. 63-75;

Aksjosemiotyka karty pocztowej II, red. P. Banaś, Wrocław 2004;

Baranowska M., *Posłaniec uczuć: prywatna historia pocztówki*, Warszawa 2003;

Duninówna H., *Warszawskie nowinki 1815-1900*, Warszawa 1970;

Furgalska A., Warchał J., *O tekstach pozdrowień*, „Socjolingwistyka” 4, red. W. Lubaś, Warszawa-Kraków-Katowice 1982, s. 123-127;

Kotłowski J., *Dawne pocztówki: historia, ikonografia, kolekcjonerstwo*, Warszawa 1998;

Kossecki J., *Podstawy nowoczesnej nauki porównawczej o cywilizacjach. Socjologia porównawcza cywilizacji*, Katowice 2003;

Немчинова Н.В., *Корпоративная поздравительная открытка: коммуникативный и деловой аспект*, „Известия Российского государственного педагогического университета имени А. И. Герцена” 2008, № 12 (86), s. 217-221;

Лингвокогнитивное описание текста корпоративной поздравительной открытки, „Вестник Кемеровского государственного университета” 2012, № 4, s. 71-75;

Низеева Д.Р., *Текстовое сопровождение поздравительных открыток (на материале немецкого языка)*, „Филология и культура” 2014, № 36, s. 41-44;

Marcjanik M., *Grzeczność w komunikacji językowej*, Warszawa 2013;

Postkarten erzählen Geschichte: Die Stadt Posen 1896-1918 // Pocztówki opowiadają historię: Miasto Poznań 1896-1918, red. S. Kemlein, Lüneburg 1997;

Paryzek P., *Poszukiwanie danych leksykalnych z tekstów elektronicznych (na materiale czasopisma naukowego)*, Poznań 2011 (praca doktorska napisana pod kierunkiem P. Wierzchońia);

20 Więcej na ten temat zob. w pracy Józefa Kosseckiego *Podstawy nowoczesnej nauki porównawczej o cywilizacjach. Socjologia porównawcza cywilizacji*, Katowice 2003.

Самбур М.В., *Открытка в контексте культуры: атрибуция, научное описание, экспонирование*, Москва 2014.

SŁOWA KLUCZE: karty pocztowe, językoznawstwo korpusowe, digitalizacja, transkrypcja, anotowanie

DANIEL DZIENISIEWICZ, PIOTR WIERZCHOŃ

THE WORK ON THE CORPUS OF POLISH AND RUSSIAN POSTCARDS OF THE SECOND HALF OF THE 20TH CENTURY. THE ISSUES OF TRANSCRIPTION AND ANNOTATION

The aim of this article is to present selected issues related to the process of creating the corpus of messages both in Polish and Russian sent via postcards in the second half of the 20th century. There will be described stages of digital processing of the postcard's text such as transcription of the text and creation of tags assigned to the following elements of the message's text, including the beginning phrase to the addressee, greetings, wishes, the final phrase and postscript. Particular points, which appear on the list of directives for people transcribing messages, will be put under thorough analysis. The reasons behind the choice of certain – often controversial – procedures will be discussed and illustrated with adequate examples. The considerations presented in the article constitute a part of the 100 000 *minihistories* project, which intends to create a full-text searchable corpus of around 10 000 of Polish and Russian post-war texts sent in postcards.

KEY WORDS: postcards, corpus linguistics, digitalization, transcription, annotating