



Badania terenowe

Beata Wójtowicz

Prezentacja chroniona prawem autorskim.



Are you ready to go to the field?!

- ▶ Zastanawiałeś się kiedyś dlaczego Twój sąsiad czasami używa angielskich słówek?
- ▶ Czy może chciałbyś zgłębić tajniki jakiegoś egzotycznego języka, który jest używany na końcu świata?
- ▶ A może zainteresowała Cię jakaś lingwistyczna teoria i masz ochotę przekonać się czy ma ona potwierdzenie w jakimś języku?

Wszystkiego możesz dowiedzieć się prowadząc badania terenowe.



- ▶ Badania terenowe to szansa na przeżycie fascynującej przygody, ale poza tym
- ▶ są niezbędne w zdobywaniu wiedzy językowej,
- ▶ dostarczają prawdziwych danych, na podstawie których można opracowywać kolejne, nowe teorie,
- ▶ zebrane dane mogą potwierdzić to, co podejrzewamy, albo dostarczyć zaskakująco nowych faktów,
- ▶ zmienić postrzeganie świata przez 'armchair linguist'.



Czym są badania terenowe?

▶ Badania prototypowe

- językoznawca pracuje z małą społecznością, daleko od domu, przez długi czas; badany język jest mało lub zupełnie nieznany, używany w naturalnym środowisku, dane są naturalne, a motywacja do badań jest *language-driven*.

▶ Badania nieprototypowe

- językoznawcy przeprowadzają wywiady pomiędzy sobą lub wśród innej dużej społeczności, blisko domu, przez krótki okres czasu; badany język jest dobrze znany a badane zagadnienia to system językowy, dane są kontrolowane a badania prowadzone pod egidą konkretnej teorii.

- ▶ Najczęściej badania terenowe posiadają cechy obu typów.



Historia

- ▶ Badania terenowe jako dyscyplinę zapoczątkował Franz Boas (pochodzący z Niemiec, 1858-1942), który wykształcił pierwsze pokolenie antropologów-językoznawców, dając podwaliny amerykańskiej szkole językoznawczej.
- ▶ Boas doszedł do przekonania, że język jest jednym z najistotniejszych przejawów danej kultury → refleksja nad wzajemnym powiązaniem języka i kultury stała się centrum obszaru badawczego językoznawstwa amerykańskiego → *anthropological linguistics, linguistic anthropology, ethnolinguistics.*



► Produkowano wtedy praktyczne opisy języków mając na uwadze ich kulturowe uwarunkowania – gramatyki, słowniki. Podkreślano związek języka z kulturą ludzi nim się posługujących i jej wpływ na morfologię, leksykę i gramatykę. Opisywano język jako manifestację kultury, psychologii i warunków życia.



Odwrót w latach '70

"Harris suggested that I undertake a systematic structural grammar of some language. I chose Hebrew, which I knew fairly well. For a time, I worked with an informant and applied methods of structural linguistics as I was then coming to understand them. The results, however, seemed to me rather dull and unsatisfying. Having no very clear idea as to how to proceed further, I abandoned these efforts and did what seemed natural; namely, I tried to construct a system of rules for generating the phonetic forms of sentences, that is, what is now called a generative grammar." (Chomsky 1975,25).



Chomsky

- ▶ Zwrot ku badaniom teoretycznym,
 - ▶ na języku macierzystym,
 - ▶ przy zupełnym porzuceniu innych języków miejscowych.
-
- ▶ Dopiero lata '90 przyniosły kolejną odsłone badań terenowych, związanych teraz głównie z badaniem języków zagrożonych.



Definicja

► Fieldwork describes the activity of a researcher systematically analyzing parts of a language other than one's native language (usually one the researcher did not speak prior to beginning fieldwork), within a community of speakers of that language, prototypically in their native land, living out their existence in the milieu and mental currency of their native culture (Everett 2001).



► Fieldwork describes the activity of a researcher systematically analyzing parts of a language, usually other than one's native language and usually within a community of speakers of that language (Sakel, Everett 2012).



Mosetén, Boliwia

- ▶ Informator
- ▶ Nagranie tekstu
- ▶ Transkrypcja → IPA / zapis przez informatora



- ▶ O czym jest tekst?
- ▶ Czy rozumiemy jakiegokolwiek słowa? →
zannotujemy je



Transkrypcja

Całość nagrania

Yäe yäerä' yoshropaiyeyak jin mi'in jae'mä sobakityi' ichhe' öchhe' öichhe' jakh Bolivia. Yäesì' ti'i nash ti'i, yäesì' ti'i chhata' ti'iyäe Fan Juan Huasna Bozo. Nä'iyäe Santa Anaya' de Mosetensi' khäki jike tse'yäe mö'yä'si'.

► Tylko tekst właściwy

Yäerä' yoshropaiyeyak in mi'in sobakityi' öichhe' jakh Bolivia. Yäesì' ti'i nash, yäesì' ti'i, chhata' ti'iyäe Juan Huasna Bozo. Nä'iyäe Santa Anaya' de Mosetensi' khäki jike tse'yäe mö'yä'si'.

► Transkrypcja końcowa

(Yäe) yäerä' yoshropaiyeyak (j)in mi'in (jae'mä) sobakityi' (ichhe' öchhe') öichhe' jakh Bolivia. Yäesì' ti'i nash (ti'i), yäesì' ti'i, chhata' ti'iyäe (Fan) Juan Huasna Bozo. Nä'iyäe Santa Anaya' de Mosetensi' khäki jike tse'yäe mö'yä'si'.



Tłumaczenie

- (1) *(Yäe) yäerä' yoshropaiyeyak (j)in mi'in (jäe'mä) sobakityi' (ichhe' öchhe') öichhe' jakh Bolivia.*
'I want to thank you, visiting this land Bolivia.'
- (2) *Yäesi' ti'i nash (ti'i), yäesi' ti'i, chhata' ti'iyäe (Fan) Juan Huasna Bozo.*
'My name is Juan Huasna Bozo.'
- (3) *Nä'iyäe Santa Anaya' de Mosetensi' khäki jike tse'yäe mö'yä'si'.*
'I was born in Santa Ana de Mosetenes because my mother was from there.'



Glosy

- (1) *Yäe yäerä' yoshropaiyeyak*
I I will thank you
(j)in mi'in (jä'mä) sobakityi'
? you uh visitor
(ichhe' öchhe') öichhe' jakh Bolivia.
in this in this in this earth Bolivia
'I want to thank you visiting this land Bolivia.'
- (2) *Yäesi' ti'i nash (ti'i) ... chhata' ti'iyäe*
my name is name well my name
Fan Juan Huasna Bozo.
Juan Juan Huasna Bozo
'My name is Juan, Juan Huasna Bozo.'
- (3) *Nä'iyäe Santa Anaya' de Mosetensi'*
I was born in Santa Ana de Mosetenes
khäki jike tse'yäe mö'yä'si'.
because then my mother from there
'I was born in Santa Ana de Mosetenes because my mother was
from there.'



Pytania

▶ Jak informator wyraża

■ zaimek 'ja'

■ relację posiadania



yäe = ja

- (1) *Yäe yäerä' yoshropaiyeyak...*
I I will thank you
'I want to thank you ...'
- (2) *Yäesi' ti'i nash (ti'i) ... chhata' ti'iyäe*
my name is name well **my** name
Fan Juan Huasna Bozo.
Juan Juan Huasna Bozo
'My name is Juan, Juan Huasna Bozo.'
- (3) *Nä'iyäe Santa Anaya' de Mosetensi'*
I was born in Santa Ana de Mosetenes
khäki jike tse'yäe mö'yä'si'.
because then **my** mother from there
'I was born in Santa Ana de Mosetenes because my mother was from there.'



▶ yäe-sĩ my

▶ ti'i-yäe my name

▶ nä'i-yäe I was born

▶ tse'-yäe my mother

▶ nä'i was born

▶ ti'i name

▶ tse' mother

▶ -yäe — posiadanie kiedy dodane do *n* lub podmiot w *v*

▶ 'my' → yäe-sĩ lub dodając -yäe do *n*



▶ -si/-si' wyraża relację posiadania 'of' – 'of me', 'of Mosetenes'

▶ -ya'/-yä' wyraża 'w' lub inne relacje przestrzenne



▶ uzyskanie dalszych wyjaśnień (*elicitation*)

■ z korpusu innych tekstów,

■ od informatora.



Początek badań

- ▶ Praca z uzyskanym tekstem.
- ▶ Praca z listą Swadesha.



Morris Swadesh (1909-1967)

Twórca metod zwanych leksykostatystyką (ustalenie stopnia pokrewieństwa języków), i glottochronologią (ustalenie czasu – dat absolutnych - rozdzielania się języków), których celem jest wyznaczanie na podstawie faktów językowych chronologii rozdzielania się języków.

Morris Swadesh (1940-50) wyszedł z założenia, że chociaż w każdym języku zasób słów (leksyka) ulega zmianom, to jednak istnieją słowa ‘stałe’, niezmiennie, potrzebne do porozumiewania się w sprawie istniejących zawsze zjawisk. Ludzie bowiem, wszędzie i zawsze muszą komunikować się co do jedzenia, picia, rodzenia się czy umierania. Według Swadesha w każdej genetycznej rodzinie językowej słowa te posiadają wspólną postać dźwiękową. Ulega ona przekształceniom fonetycznym w miarę rozdzielania się i rozwoju języków, ale są to przekształcenia dotyczące takiej formy, która jest wspólna dla całej rodziny językowej.



Założeniem teorii jest, że liczba zmian fonetycznych wewnątrz tego podstawowego zasobu słów jest wymierną statystycznie stałą liczbową. Oznacza to, że liczba zmian fonetycznych w tym zasobie jest proporcjonalna do ilości czasu, jaki upłynął od oddzielenia się języków od wspólnego, wyjściowego przodka.



- ▶ Szybkość zmian w zakresie podstawowego słownictwa obserwowana w długich okresach czasu jest mniej więcej jednakowa, mimo iż w krótkich odcinkach może być zróżnicowana.
- ▶ Z badań wynika, że w ciągu ok. tysiąca lat społeczność językowa zachowuje ok. 81% niezmiennych jednostek podstawowego słownictwa.
- ▶ Jeśli następuje rozpad i dwa języki rozwijają się oddzielnie, ta wielkość wynosi 66% (część wspólna w obu językach).



- ▶ Jaka jest definicja 'słownictwa podstawowego'?
- ▶ Zbadano dotychczas zbyt mało języków, aby można było wyciągnąć wnioski o charakterze uniwersalnym.



► Amerykański językoznawca, Morris Swadesh, wyszedł z założenia, że uniwersalne pojęcia wynikają z uniwersalnych warunków życia i etapów życia każdego człowieka. Każdy zna, z własnego doświadczenia, słońce, księżyc i gwiazdy, noc i dzień, wiatr i deszcz, wodę i ogień, itd. Każdy zna też ciało ludzkie. -> Właśnie wśród tego typu pojęć należy szukać wyrazów, które mają odpowiedniki we wszystkich językach.



Lista Swadesha

► Lista podstawowego słownictwa – oryginalnie 207 słów zredukowane potem do 100:

I, you, he, we, you (pl), they, this, that, here, there, who, what, where, when, how, not, all, many, some, few, other, one, two, three, four, five, big, long, wide, thick, heavy, small, short, narrow, thin, woman, man, man (human being), child, wife, husband, mother, father, animal, fish, ... , hair, head, ear, eye, nose, mouth, tooth, ..., smell, fear, sleep, live, die, kill, fight, black, night, day, year, warm, cold, full, ...



Zapożyczenia

- ▶ Pola semantyczne najbardziej podatne na zapożyczenia to religia i wierzenia, ubrania i wyrazy ze sfery gospodarstwa domowego, które pochodzą z dziedzin najbardziej podatnych na kontakt międzykulturowy.
- ▶ Najbardziej stałe grupy słownictwa, najmniej podatne na zapożyczenia, to terminy pokrewieństwa, nazwy części ciała, relacje przestrzenne i zmysły; ponadto, do tej grupy zaliczamy też wyrazy funkcyjne, czy zaimki pytajne.
- ▶ W grupie wyrazów zapożyczonych częściej odnotowuje się rzeczowniki konkretne niż abstrakcyjne.

(por. <https://doi.org/10.1515/9783110591484>)



Badania jedno- czy dwujęzyczne?

- ▶ Dwujęzyczne – przy pomocy lingua franca
 - używamy innego niż badany język porozumiewając się z badanymi osobami,
 - ponieważ nie znamy badanego języka,
 - ułatwia komunikację ale może mieć wpływ na dane → badani przenoszą konstrukcje z lingua franca na swój język,
 - znacznie przyspiesza badania – ważne w sytuacji ograniczonego czasu,
 - ułatwia ustalenie formalności – umówienie się, płatność.



Badania jedno- czy dwujęzyczne?

- ▶ Jednojęzyczne – tylko w badanym języku, gdy
 - badani nie znają innego języka,
 - zależy nam na nauczaniu się badanego języka (pożądane jeśli piszemy gramatykę języka),
 - widzimy wyższość tej metody nad dwujęzyczną → chłoniemy każdy dźwięk, gest, sytuację,
 - chcemy okazać większy szacunek badanym, poprzez pominięcie np. języków dominacji (które mogłyby być naszymi lingua franca),
 - nie zawsze mamy możliwość przetłumaczenia poszczególnych zwrotów/wyrażeń na inny język.



Typy danych

- ▶ Dane pisane – bardziej złożone pod względem struktury; łatwiejsze do zebrania.
- ▶ Dane mówione – bardziej spontaniczne, innowacyjne; trudniejsze do pozyskania.



Typy języków

- ▶ Duże języki – wielu użytkowników, łatwy dostęp, dostępne materiały o języku i w języku.
- ▶ 'Małe' języki, języki zagrożone – mało użytkowników?, trudniejszy dostęp?, mało lub brak materiałów o i w języku, brak użytkowników biegłych w języku, niemożność zebrania danych z różnych rejestrów języka, trudności z kontynuowaniem badań (względy personalne, śmierć ostatniego użytkownika) ALE odkrywanie nieznanego, tworzenie pierwszego opisu/gramatyki/słownika.



Konflikt interesów

- ▶ W trakcie badań mogą pojawić się problemy z
 - Informatorem/nauczycielem,
 - Społecznością użytkowników języka,
 - Organizacjami pracującymi na danym terenie,
 - Innymi językoznawcami.



Typy metod badawczych

► Jakościowe

- praca na danych naturalnych zebranych wśród małej liczby użytkowników,
- różne typy danych – opowiadania, konwersacje, wywiady,
- zebrana próbka danych jest niewielka,
- nacisk położony jest na językowe zachowanie jednostki raczej niż na porównanie zachowań pomiędzy różnymi użytkownikami,
- zwykle metoda wykorzystywana w badaniach prototypowych.



Typy metod badawczych

- ▶ Ilościowe – wykorzystywane w badaniu zmiennych wśród wielu użytkowników;
 - często wykorzystuje metodę manipulacji (jedna grupa zostaje poddana jakiejś manipulacji a druga nie);
 - dane sprowadzane są do liczb (np. wymowa 'r' wśród danej populacji przedstawiona procentowo);
 - potrzeba dużo większej ilości danych (generalizacja wyników);
 - dane pozyskiwane z kwestionariuszy, analizowane statystycznie – badanie może zostać powtórzone
→ uogólnione sądy na temat języka i zachowań użytkowników.



Typy metod badawczych

► Mieszane

- wykorzystuje oba typy danych – ilościowe i jakościowe – np. zbiór danych ilościowych w oparciu o kwestionariusz + pytania otwarte, do analizy jakościowej.



Metadane

- ▶ Kod – gatunek tekstu – narracja, konwersacja
- ▶ Numer pliku – gdzie go znaleźć w bazie
- ▶ Temat – o czym jest tekst
- ▶ Daty – kiedy tekst został nagrany, kiedy transkrybowany
- ▶ Informacja o informatorze – kto to, skąd, płeć, wiek, stan cywilny, zawód, znajomość języków
- ▶ Zespół badawczy – kto pracował przy projekcie
- ▶ Miejsce – gdzie tekst nagrano
- ▶ Inne

Zbieranie danych tekstowych

- ▶ Narracja – opowiadania, zarówno teksty literatury ustnej, jak i nieformalne wypowiedzi typu 'co robiłem w wakacje'.
- ▶ Przepisy – informator podaje przepis na coś, np. jak zrobić łuk (→ tryb rozkazujący, zależności czasowe).
- ▶ Objasnienia – wyjaśnienie np. zjawisk kulturowych, jak zostać szamanem (→ aspekt, czas).
- ▶ Zagrzewające do działania.
- ▶ Konwersacje – najważniejszy i zarazem najtrudniejszy do zebrania typ tekstu.



Nagranie tekstu



Transkrypcja



Tłumaczenie całego tekstu



Tłumaczenie słowo po słowie/ morfem po morfemie
(glosy)



Objaśnienie



Transkrypcja

- ▶ 1h nagrania : 4h spisywania
- ▶ Użycie ortografii języka (jeśli istnieje, jeśli znamy lub ktoś może to zrobić dla nas)
- ▶ Użycie IPA
- ▶ Przy pomocy programów komputerowych
 - [Transcriber](#) (łączy nagranie audio z tekstem)
 - [ELAN](#) (łączy nagranie audio/video z tekstem, możliwość tłumaczenia i glosów)
 - [CLAN](#) (łączy nagranie audio/video z tekstem, tłumaczenie i glosy + anotacja i statystyczna analiza danych) [video](#)



Tłumaczenie

- ▶ Ogólne tłumaczenie całego tekstu.
- ▶ Tłumaczenie poszczególnych zdań.
- ▶ Tłumaczenie słowo po słowie.
- Tłumaczenie słowo po słowie nie zawsze równa się tłumaczeniu całego sensu zdania.

Ampigibongo

Uderzył mózg

'Zastanawiał się'



Glosy

- ▶ Niezbędne w analizie języka.
- ▶ Tłumaczenie morfemów (kolejny poziom). Metoda ta umożliwia zrozumienie i analizę przykładów z języków, których nie znamy.

• *Mitume wawili walianguka*

Mi-tume

wa-wili

wa-li-anguka

KL.4-apostoł

KL.2-dwa

KL.2-PAST-przewrócić się

'dwóch apostołów przewróciło się'

■ *Nilimwona mtoto*

Ni-li-mw-ona

m-toto

SBJ.1.SG-PST-OBJ.3.SG-widzieć

KL.1-dziecko

'widziałam dziecko'



► Próba standaryzacji

■ Leipzig glossing rules

■ <http://www.eva.mpg.de/lingua/resources/glossing-rules.php>

► Programy komputerowe

■ ELAN

■ CLAN

■ [FieldWorks](#) - [Quick tour](#)



Elicitation - objaśnienie

- ▶ Wywiad z użytkownikiem/kami języka
- ▶ Uzyskanie dodatkowych informacji o języku, wyjaśnienie niezrozumiałych danych, potwierdzenie wyników, ...



Wyniki badań

- ▶ Publikacja zebranych materiałów
- Korpus tekstów (mówionych <> pisanych)
- Ogólnodostępny?
- metadane!
- Anotowany (FieldWorks)
- Publikacja zbioru opowiadań/ tekstów
- Archiwizacja (wszystkich materiałów!) na platformie internetowej
- [Open Language Archives Community OLAC](#)
- [Documentation of Endangered Languages DoBeS](#)
- Publikacja bazy danych w Internecie



Prezentacja badań

► Publikacje, konferencje →

- prezentacja metodologii badań

- w prezentacji badań metodą ilościową ważna jest możliwość powtórzenia badania



Place and participants of the study

- Who are your speakers and teachers?
 - Who are the researcher(s)?
 - How did you find your speakers and teachers? Did you give any consideration as to general practicality, ease of access, suitability for the study, age-range, other sociolinguistic factors, languages?
 - Where did you carry out your study? What were your considerations in choosing a certain place or setup for the study?
-

Type of fieldwork methods

- Which fieldwork methods did you use (quantitative, qualitative, as well as details, e.g. elicitation and text analysis)?
-

Materials for the empirical investigation

- Which types of materials did you use and how did you prepare them (e.g. how did you build up your questionnaire / design your experiments)?
 - What recording equipment did you use?
 - What setup in the recording room did you use?
-

Mode of carrying out the data collection

- *Questionnaires*: Did you stay with the participants while they filled in the questionnaires independently? Did you give the speakers 'background information' about your study that could have influenced the results (e.g. positive attitudes)? Did the speakers know exactly what it was you were after or did you pose a general question?
 - *Elicitation and text collection*: How did you carry out your recordings? Did you take part in your recording (as an active participant)? Did you (attempt to) trigger a certain language behaviour (e.g. code-switching) during or before the recording?
 - *Experiments*: What was the setup of your experiment(s)?
-

Analysis of the raw data

- How did you prepare your raw data for analysis?
 - Did you use a statistics program?
 - Did you use a specific computer program (e.g. Transcriber, ELAN)?
 - Did you transcribe and analyse all your recordings or only parts (how many minutes did you record altogether)?
 - How did you transcribe and translate your data (with the help of a speaker?)
 - How did you compare the results from your questionnaire study?
-

Ethics

- What were your general ethical considerations?
 - Did you record the names of teachers – did they agree to this?
 - Did you promise your teachers anonymity?
 - Did you pay your speakers or teachers?
-



Wynik badań to Gramatyka języka

► Możemy oprzeć się na wzorcu

■ Lingua Descriptive Studies Questionnaire (N. Smith, B. Comrie)

► Potrzebny duży korpus tekstów – reprezentatywny.

■ Zawiera wszystkie elementy skończonych zbiorów (np. wszystkie zaimki),

■ Nie ma 'dziur' w danych potrzebnych do analizy (np. brak przykładu na wystąpienie danego dźwięku),

■ Jest kilka przykładów wystąpień danej jednostki (konstrukcji, sufiksu),

■ Teksty są przydatne również dla innych dyscyplin,

■ Każdy kolejno dodany materiał zawiera struktury i znaczenia, które już można znaleźć w korpusie.



Wynik badań to Słownik

- ▶ Cel → grupa odbiorców
- opracowanie ortografii i formy standard języka
- dokumentacja języka
- dokumentacja językoznawcza (hasła z b. rozbudowaną informacją gramatyczną)
- słownik dwujęzyczny

- ▶ Lista haseł
- ▶ Korpus tekstów



Znaczenie badań

► [https://www.youtube.com/watch?v=1YFn20iV4O](https://www.youtube.com/watch?v=1YFn20iV4OA)
A