

Wstęp

Ten artykuł to niezbędny przewodnik dla każdego, kto chce poważnie i efektywnie współpracować przy projekcie Tatoeba. Przewodnik jest dość długi, ale mamy nadzieję, że zrozumiały.

1. Musisz rozumieć kontekst projektu

Kiedyś pewnie napiszę bardziej szczegółową historię, tutaj tylko podstawowe fakty, o których trzeba pamiętać.

- Z projektem ruszyłam w roku 2006. Motywacją do podjęcia go była pasja do języków i moja frustracja, że nie ma w Sieci odpowiedniego słownika.
- Projekt skupia się na **zdaniach** i na nie kładę nacisk. A to dlatego, że uważam, że zdań z przykładami było (i wciąż jest) zdecydowanie za mało. Jeśli chcesz współpracować, proszę o dodawanie tylko **całych zdań**.
- Przez jakiś czas ciągnęłam projekt praktycznie sama. Dopiero trzy lata później, w roku 2009, dołączyli do mnie pozostali (wszyscy studenci informatyki), by pomóc mi dopisać więcej funkcjonalności.
- Tatoeba to projekt **niekomercyjny**. Nie jesteśmy firmą, nie otrzymujemy pieniędzy za naszą pracę. Robimy to w naszym **czasie wolnym**.
- Szczerze mówiąc, nie wykluczamy możliwości komercjalizacji z przyszłości, ale tylko wtedy, jeśli wypracujemy innowacyjny, spójny i etyczny model biznesowy (oby!). Wykluczamy możliwość napakowania strony reklamami, podbijania ruchu albo zmuszania użytkowników do płacenia za dostęp do danych.

2. Jak wygląda struktura korpusu?

To nie jest takie banalne, postaram się wyjaśnić to tak, żeby wszyscy zrozumieli.

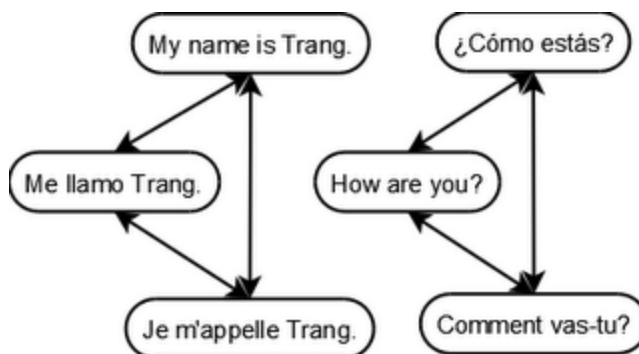
Korpus nie jest zbudowany jako **tabela**, ale jako **graf**. Co to oznacza? Wyobraźmy sobie, że masz wziąć fragment korpusu i zapisać go na kartce. Pewnie powstanie coś takiego:

Angielski	Francuski	Hiszpański
My name is Trang.	Je m'appelle Trang.	Me llamo Trang.
How are you?	Comment vas-tu?	¿Cómo estás?

...	...	...
-----	-----	-----

To jest struktura **tabelaryczna**. Są tu **wiersze** i **kolumny**: wiersz zawiera zdania o tym samym znaczeniu, a kolumna - zdania w tym samym języku. To pierwsze i oczywiste rozwiązanie, ale nasz korpus zbudowany jest inaczej.

Struktura naszego korpusu jest następująca:



To struktura **grafu**. Są tu **węzły** i **strzałki**: każdy węzeł reprezentuje zdanie, a każda strzałka - połączenia między dwoma zdaniem. Kiedy dwa zdania są połączone, mają to samo znaczenie.

Sposób dodawania nowej treści bardzo różni się w zależności od struktury. Istotny wniosek polega na tym, że **możesz dodawać wielokrotne tłumaczenie w tym samym języku** do tego samego zdania. Uważasz, że istnieje więcej niż jedna możliwość przetłumaczenia zdania, ale nie jesteś pewien, która jest najlepsza? Dodaj wszystkie!

Jakie są inne implikacje takiej struktury? Jedźmy dalej.

3. Nie zwracaj uwagi na inne tłumaczenia

Kiedy tłumaczysz zdanie, w istocie **dodajesz zdanie** (węzeł) i **połączenie** (strzałkę) między oryginałem a swoim tłumaczeniem. A więc jedyne, co powinno cię obchodzić, to sformułowanie poprawnego przekładu zdania głównego (będzie na samej górze, większą czcionką).

Konkretnie, gdybyś był w takiej sytuacji i miał dodać tłumaczenie hiszpańskie do zdania angielskiego:

How are you?
=> *Comment vas-tu?*

to możesz dodać "*¿Cómo estás?*" (czyli "Jak się masz?" - do znajomego), ale również "*¿Cómo está usted?*" (czyli "Jak się pan miewa?" - do kogoś, z kim jesteśmy na pan/pani). Ale możesz dodać oba zdania (bo można dodawać kilka tłumaczeń w tym samym języku).

Jeśli znasz francuski, **nie ma znaczenia**, że tłumaczenie zdania w tym języku jest sformułowane w formie odpowiedniej tylko dla zwracania się do przyjaciół (czyli na "ty"). Zważaj tylko na to, by twoje tłumaczenie było poprawnym ekwiwalentem **zdania wyjściowego**, czyli w tym wypadku angielskiego. Poprawne tłumaczenie oznacza, że gdyby ktoś kiedyś chciał przetłumaczyć twoje zdanie z powrotem na angielski, to zdanie "*How are you?*" będzie również poprawną wersją.

4. Nie tłumacz dosłownie!

Nie potrzeba nam zdań, które brzmią jak tłumaczone przez robota. Chcemy takich, które faktycznie pojawiłyby się w ustach rodzimego użytkownika danego języka (native speaker). Tłumaczenie nie jest taką łatwą sztuką, wiadomo. Jeśli tłumaczysz coś na swój język ojczysty, to zawsze, powtarzam, zawsze przeczytaj jeszcze raz swoje tłumaczenie jako niezależne zdanie. Zadaj sobie pytanie, czy tak rzeczywiście mówią twoi rodacy. Gdybyś z jakichś powodów musiał się zdecydować na tłumaczenie dosłowne, możesz użyć komentarza i poinformować o tym.

Jeśli jednak nie tłumaczysz na swój język ojczysty (do czego masz prawo), twoje zdania mogą nie brzmieć jak zdania wypowiedziane przez rodzimych użytkowników. W końcu to projekt społecznościowy, stąd mamy nadzieję, że w jakimś momencie jakiś rodzimy użytkownik natrafi na takie wątpliwe zdanie i poprawi, jeśli będzie ono brzmieć w jego języku nienaturalnie.

Najważniejsze jest zrozumienie, że w Tatoeba nie chodzi tylko o dostarczanie tłumaczeń, ale i gromadzenie danych na temat języka. Można by zredukować Tatoeba do kompletowania zdań bez ich tłumaczenia. Gdybyśmy mieli wyciągnąć z bazy np. tylko zdania w języku włoskim, to wszystkie takie zdania powinny być dobrą reprezentacją tego języka.

Zdania to warstwa podstawowa. Kolejna warstwa to powiązania między zdaniami, ale korpus powinien mieć sens również bez takich powiązań.

5. Nie edytuj zdania, jeśli samo w sobie jest poprawne

Jak napisałam wyżej, Tatoeba można rozumieć jako narzędzie do prostego dodawania zdań bez ich tłumaczenia. A więc, zanim zmodyfikujesz jakieś zdanie, przypatrz się mu, bez zwracania uwagi na inne jego tłumaczenia, zadając sobie pytanie: "Czy to zdanie zawiera błędy

ortograficzne czy gramatyczne? A może dziwnie brzmi?” Jeżeli odpowiedź brzmi “nie”, nie edytuj go, **zostaw jak jest**.

To zastrzeżenie jest potrzebne, bo możesz czuć pokusę zmiany zdania w taki sposób, by znaczenie pasowało do pozostałych zdań.

Przyczyną może być na przykład chęć zmiany zdania w kierunku bardziej dosłownym. Ale nie tędy droga. Przecież jeśli nie chcemy tłumaczeń dosłownych (patrz zasada numer 4), to również nie potrzebujemy zmieniania zdań na bardziej dosłowne!

Przyczyna może być jeszcze inna - mianowicie że zdania w ogóle do siebie nie pasują. Weźmy przykład:

My name is Trang.

=> *Je m'appelle Trang.*

=> *Vamos a la playa.*

Jak można zauważyć, zdanie hiszpańskie (które znaczy "Chodźmy na plażę") nie ma nic wspólnego ze zdaniem angielskim.

Może niespecjalnie czujesz się mocny w hiszpańskim, więc zechcesz zmienić zdanie angielskie. I tu rodzi się problem: co w takim razie ze zdaniem francuskim? Teraz z kolei ono przestanie pasować do angielskiego...

A może jesteś rodzimym użytkownikiem hiszpańskiego i zechcesz zmienić zdanie hiszpańskie? W tym konkretnym przypadku jest to dopuszczalne, ponieważ zdanie hiszpańskie nie jest powiązane z żadnym innym. Jednak jeśli zdanie hiszpańskie zostało z kolei przetłumaczone np. na włoski, “poprawienie” zdania hiszpańskiego zrodzi konflikt między nim a tłumaczeniem na włoski.

Poza tym jest jeszcze jeden problem, o którym mało kto pamięta - kiedy zmieniasz znaczenie zdania, możesz usunąć jakieś unikalne słowo czy wyrażenie. A jeśli hiszpańskie zdanie będzie jedynym zdaniem w bazie z wyrazem “playa”?

Co robić zatem w takiej sytuacji? Najlepiej dodaj nowe, poprawne tłumaczenie na hiszpański (*Me llamo Trang*) i odlinkuj aktualne tłumaczenie. Uwaga: nie każdy może odlinkowywać. Takie uprawnienia mają tylko “zaufani użytkownicy” (“trusted users”). Możesz jednak zamieścić komentarz (najlepiej po angielsku) z uwagą, by odlinkować dane zdanie.

6. Nie zmieniaj języka zdań

Jeśli flaga przy zdaniu jest błędna (np. pojawia się flaga chińska, a zdanie jest po japońsku), to oczywiście można to zmienić, ale nie o tym tu mówimy.

Chodzi o to, by nie zastępować tego zdania po japońsku zdaniem chińskim o tym samym znaczeniu (co oczywiście dotyczy wszystkich języków). To nie powinno zdarzać się często, ale jeśli się tak zdarzy, nie poprawiaj w taki sposób.

Problem polega na tym, że zdanie może być powiązane z danymi, które zależą od języka, w jakim jest sformułowane. Na przykład komentarze: użytkownicy mogą zamieszczać komentarze przy zdaniach, a te będą ważne tylko w przypadku oznakowania zdania określonym językiem.

W chwili obecnej problem dotyczy bardziej zdań japońskich, do których przypisane są określone adnotacje. One nie są wyświetlane, bo zwykłym użytkownikom są niepotrzebne. Ale jeśli zamienisz np. zdanie japońskie na angielskie, to adnotacje stracą jakikolwiek sens.

7. Dodawaj komentarze do właściwego zdania

Kiedy zamieszczasz komentarz, będzie on powiązany tylko ze zdaniem głównym, więc upewnij się, do jakiego konkretnie zdania on się odnosi. Na przykład, chcesz zwrócić uwagę na błąd ortograficzny, jak tu:

My name is Trang.
=> *Je m'appel Trang.*
=> *Me llamo Trang.*

Jak widać, zdanie francuskie zawiera błąd. Powinno być “appelle”, a nie “appel”. Jeśli jednak zamieścisz komentarz w tym miejscu, będzie on powiązany ze zdaniem angielskim (bo to ono jest na górze, więc jest zdaniem głównym). A nie tego przecież chcemy. Kliknij najpierw na zdanie francuskie. Układ zmieni się na:

Je m'appel Trang.
=> *My name is Trang.*
=> *Me llamo Trang.*

Teraz dopiero dodaj komentarz.

A co w sytuacji, kiedy chcesz zwrócić uwagę na błędne tłumaczenie? Komentarz będzie się odnosić do więcej niż jednego zdania, więc gdzie go zamieścić? W idealnym świecie powinna być możliwość dopisania komentarza do **powiązania** między dwoma zdaniami. Takiej możliwości jednak nie mamy, mamy tylko możliwość komentowania zdań. W takiej sytuacji sam musisz zdecydować, gdzie zamieścić komentarz. Pamiętaj tylko, że będzie on ważny tylko

wtedy, kiedy będzie odnosić się do głównego zdania.

8. Nie dodawaj zdań ze źródeł chronionych prawem autorskim

Korpus jest dystrybuowany na licencji [Creative Commons Attribution](#) (CC-BY). Umożliwia ona każdemu dalsze wykorzystanie danych w dowolny sposób, pod warunkiem odesłania do Tatoeba w dziele pochodnym.

Jako współuczestnik projektu wyraziłeś zgodę na warunki użytkowania (których oczywiście nie przeczytałeś), a więc swoje tłumaczenie również udostępniasz na zasadach licencji CC-BY. Oznacza to, że możemy przetwarzać dane wprowadzone przez użytkowników pod warunkiem dodania adnotacji o autorstwie. Dlatego waszą pracę przetwarzamy w ramach projektu Tatoeba, a współautorów opisujemy poprzez logi i statystyki.

Zamieszczanie zdań w ramach CC-BY oznacza, że ponosisz również pewną odpowiedzialność w zakresie wprowadzanych danych. Musisz pamiętać, że **nie masz prawa** legalnie rozprowadzać danych, jeśli zostały skopiowane ze źródła, które nie stwierdza wyraźnie, że masz do tego prawo. Na przykład nie możesz (legalnie) skopiować wszystkich zdań z podręcznika i wrzucić na Tatoeba.

Bez obaw, ani ty, ani my nie pójdziemy do więzienia ani nie zostaniemy zarżnięci kosmicznym odszkodowaniem do końca życia, jeśli wrzucisz kilka zdań z jakiegoś podręcznika (przynajmniej mamy taką nadzieję...). Jednak prawo zakazuje nam brania sobie bez pozwolenia czyjejś pracy i wykorzystywania w naszym projekcie. Tworzenie i tłumaczenie zdań stanowi pracę, więc prosimy zwracać uwagę na to, skąd bierzecie swoje zdania. Najlepiej twórzcie własne zdania lub czerpcie z książek znajdujących się w domenie publicznej.

Jeśli dodałeś lub zobaczyłeś zdania skopiowane ze źródła chronionego prawem autorskim, zmień w nich kilka słów, by nie były to te same zdania. Ewentualnie przekonaj ich autorów, by zgodzili się opublikować swoje dzieła w ramach licencji CC-BY, byśmy mogli ich swobodnie użyć.

Nie będę się upierała, że to wszystko ma sens (sama oczywiście uważam, że nie), ale będziemy wdzięczni, jeśli każdy zrobi, co do niego należy, by nas nie pozwano.

9. Nie zamieszczaj w zdaniach żadnych adnotacji

Chcemy, by zdania były maksymalnie “surowe” i naturalne, a więc nie zawierały żadnych adnotacji. **Nie chcemy** zatem zdań typu :

1. I (female) am happy.

2. It's raining cats and dogs. (idiom)
3. I like her/him.

Co do zdań numer 1 i 2, to jeśli czujesz potrzebę zaznaczenia, że zdanie jest przysłowiem, idiomem czy jest w formie żeńskiej itp., to zamieść komentarz (lub tag, jeśli jesteś zaufanym użytkownikiem), ale nie zamieszczaj tego typu informacji bezpośrednio w zdaniu.

Co do zdania numer 3, to zamiast komasowania dwóch form w jednym, możesz rozdzielić je na dwa niezależne. Pamiętaj, masz prawo dodawać wielokrotne tłumaczenia tego samego zdania. Możliwa jest więc sytuacja następująca:

Je l'aime bien.

=> *I like her.*

=> *I like him.*

Adnotacji nie chcemy z kilku powodów:

1. Mogą stanowić problem w przypadku użycia naszych danych w celu np. udoskonalania systemów przetwarzania języka naturalnego.
2. Twoje tłumaczenie może z kolei zostać przetłumaczone na inny język, wtedy innym będzie trudniej radzić sobie ze zdaniami, które zawierają alternatywy (typu "him/her").
3. Jeśli będziemy chcieli nagrać wersję audio danego zdania, będziemy musieli wybrać, którą wersję zarejestrować, adnotacje w tym nie pomagają.

10. Nie bój się pisać do nas

Zdajemy sobie sprawę, że projektowi Tatoeba daleko do doskonałości, więc zapraszamy do kontaktu - [powiedz nam](#), co sądzisz, napisz, czego brakuje (ale sprawdź, czy nie było już o tym mowy na forum na [Tablicy ogłoszeń](#)). Powiedz nam też o jakichkolwiek zauważonych błędach, niejasnych objaśnieniach czy napotkanych usterkach. Pamiętaj tylko, że nie mówimy po polsku ;-)

Wiemy poza tym, że Tatoeba to fantastyczny projekt, więc nie krępuj się i powiedz nam, że też ci się podoba :P

11. Jeśli sam umiesz programować, nie czekaj na nas

Kontakt jest mile widziany, ale jeszcze bardziej oczekujemy **inicjatywy**. Jest tyle fajnych rzeczy do zrobienia, że nie możemy zajmować się wszystkim.

Na przykład: my rozprawdzamy **cały korpus**, ale są tacy, co raczej nie potrzebują **wszystkich** zdań we **wszystkich** językach. Możesz na przykład potrzebować tylko zdań polsko-angielskich. Wtedy zamiast prosić nas i czekać, aż przygotujemy plik ze zdaniami polsko-angielskimi,

możesz sam napisać skrypt wyciągający z plików bazy tylko to, czego potrzebujesz. W takim wypadku bardzo prosimy o informację.

To tylko przykład, ale jeśli jesteś programistą, wiele rzeczy możesz zrobić sam, nie czekając, aż zrobimy to my. Oczywiście poinformuj nas o tym, byśmy nie planowali roboty nad czymś, co robi już ktoś inny.

Pamiętaj też, że w sumie jesteśmy oparci na open source (na licencji AGPL), ale tego aspektu właściwie nie “promujemy”, ponieważ:

1. Kod jeszcze nie spełnia naszych standardów elegancji... Wciąż jeszcze wiele jego fragmentów przyprawia mnie o mdłości
2. Wciąż jeszcze nie mamy solidnej metodologii i organizacji naszej pracy, a ja nie mam czasu na zarządzanie większą grupą ludzi

Jeśli jednak podoba ci się nasz projekt i chcesz dołączyć do ekipy projektowej, napisz do nas =)

12. Zaznacz swój język w profilu

Dla tych, co jeszcze nie wiedzą: możesz edytować swój profil, klikając na swój nick (na górze, na belce menu).

Ponieważ Tatoeba dotyczy języków, bardzo pomocne dla innych użytkowników będzie dowiedzenie się, jakimi językami władasz, w jakim stopniu, oraz, oczywiście, jaki jest twój język ojczysty. Nie ma tam wprowadzić określonego pola pod nazwą “języki”, ale możesz napisać o tym sam w polu “opis” (sekcja “Coś o sobie”).

Przypomnij również innym użytkownikom, by ujawnili znane im języki, szczególnie jeśli są autorami tłumaczeń.

13. Zachęcaj i poprawiaj nowych (i nie całkiem nowych też!) użytkowników

W projekcie takim jak Tatoeba społeczność jest rzeczą niezmiernie istotną - nie osiągniemy ambitnych celów bez silnej społeczności. Jak zbudować silną społeczność? Po pierwsze, nie sprawiamy, by nowi czuli się zagubieni i izolowani.

Częściowo zależy to od samego systemu. Musi być zaprojektowany tak, by nie tylko umożliwiał, ale i zachęcał użytkowników do interakcji. Tatoeba nie jest jeszcze pod tym względem idealna, ale mamy tu narzędzia, które to umożliwiają (wiadomości prywatne, tablicę ogłoszeń, komentarze).

Częściowo zależy to jednak również od samej społeczności - ona sama musi mieć wolę integracji i rozbudowy. Jeśli więc ktoś zadaje pytanie, na które umiesz odpowiedzieć, zrób to. A jeśli zauważysz, że ktoś się myli, powiedz mu o tym. Jeśli widzisz, że ktoś jest aktywnym użytkownikiem, miło mu będzie, jeśli wyślesz parę słów uznania (publicznie, na tablicy, albo prywatnie).

Ogólnie rzecz biorąc, jeśli masz pomysł, jak sprawić, żeby Tatoeba była jeszcze przyjemniejszym miejscem interakcji użytkowników, to wal śmiało!

14. Nieś słowo w świat!

I na koniec: tobie się nasz projekt podoba, nam się podoba, wszyscy chcemy, by stał się najbardziej rewelacyjnym narzędziem językowym w historii, więc ściągaj nowych użytkowników!

Przecież uczestniczyć w nim może każdy, kto tylko umie czytać i pisać. Nie trzeba być poliglotą. Jeśli “tylko” będziesz przeglądać zdania w poszukiwaniu błędów, poprawiać je i wskazywać usterki, to i tak będzie to wielka pomoc. Im więcej będzie ludzi, im więcej błędów wytropimy, tym nasza baza będzie lepsza i bardziej wiarygodna. Wszyscy żyć będą długo i szczęśliwie :-)

15. Rozmawiaj z innymi, ale międzynarodowo

Pamiętaj: Tatoeba ma wprowadzić tłumaczenie interfejsu oraz paru innych dokumentów na język polski, ale nie jest projektem polskim. Wszelka korespondencja z adminami musi się odbywać w językach, które znają (angielski, francuski). W ogóle zamieszczanie komentarzy i postów na forum, które mają trafić do wszystkich użytkowników, nie tylko do Polaków, powinno odbywać się w języku bardziej międzynarodowym niż polski.

16. Jak tłumaczyć na polski?

Polszczyzna jest jednym z języków mocno reprezentowanych w projekcie Tatoeba. Polscy użytkownicy dbają o to, by język tłumaczeń był językiem poprawnym, naturalnym i literackim. Co to oznacza?

- **Ortografia i poprawność.** Jeśli nie jesteś pewien, czy wiesz, jak się pisze dane słowo, sprawdź je, zanim klikniesz “dodaj tłumaczenie”. Podobnie, jeśli masz wątpliwość gramatyczną. Zdanie z błędami ortograficznymi i gramatycznymi jest bezużyteczne i wprowadza tylko zamęt, zamiast wzbogacać projekt.

- **Naturalność.** Przed wprowadzeniem tłumaczenia zadaj sobie pytanie, czy stworzone przez ciebie zdanie jest normalnym, naturalnym zdaniem polskim. Czy tak powiedzą Polacy? Jeśli masz wątpliwości, przemyśl ponownie, przeredaguj je.
- **Język ogólny.** Jak można się zorientować, język tłumaczeń należy do ogólnego rejestru literackiego języka polskiego - nie szalej nadmiernie z językiem potocznym (aczkolwiek pewne elementy oczywiście mogą się znaleźć, jeśli jest to uzasadnione), slangiem, dialektami, archaizacją ani urzędową sztywnością. Wulgaryzmów używajmy wyłącznie wtedy, kiedy tłumaczone zdanie ewidentnie taki zawiera.

Parę spraw szczegółowych:

- W języku polskim używamy **liter z diakrytykami** (ogonkami), czyli ą, ę, ć, ł, ó,ń, ż, ź. Zdania bez koniecznych "ogonków" są błędne i muszą albo zostać poprawione, albo będą skasowane!
- Najczęściej znanymi językami, z których pewnie będziecie tłumaczyć, są takie języki jak angielski, francuski, rosyjski. Pamiętaj, że każdy język obcy ma swoją specyfikę, której nie możemy kopiować dosłownie do zdania polskiego! Na przykład angielszczyzna lubi zaimki dzierżawcze. Angielskiego zdania *He scratched his head with his hand* **nie tłumaczymy** na polski jako *On podrapał się jego ręką po jego głowie*, ani nawet *Podrapał się swoją ręką po swojej głowie*, tylko jako *Podrapał się ręką po głowie*. Jeśli w pierwszej chwili takie zdanie napisałeś, przeczytaj je jeszcze raz, zastanów się nad jego naturalnością i wyrzuć wszelkie zbędne zamki.
- W większości języków europejskich zaimków osobowych używa się znacznie częściej niż w polszczyźnie. Polski jest językiem tzw. pro-drop, co oznacza, że zwykle zaimki są domyślne, i znacznie naturalniejsza jest wypowiedź *Wstał i wyszedł* aniżeli *On wstał i wyszedł*. Spójrz na przetłumaczone przez ciebie zdanie i wyrzuć wszystkie niekonieczne zaimki.
- Zaimki drugiej osoby zasadniczo piszemy wielką literą **tylko wtedy, kiedy zwracamy się do konkretnej osoby, znanej nam z imienia i nazwiska** - np. w liście albo w mailu. W zdaniu przykładowym w Tatoeba, którego przecież nie kierujemy do jakiegoś konkretnego odbiorcy, nie ma powodu używać wielkich liter.
- Zdania w innych językach mogą być tłumaczone na polski w różny sposób, w zależności od adresata, do którego się zwracamy, ponieważ polszczyzna jest językiem stosunkowo precyzyjnym jeśli chodzi o wykładniki osobowości. Np. angielskie zdanie *Do you often go there?* może być odpowiednikiem całego mnóstwa polskich zdań, jak choćby *Czy często tam chodzisz?* - *Czy często tam chodzicie?* - *Czy pan często tam chodzi?* - *Czy panowie często tam chodzą?* - *Czy panie często tam chodzą?* itp. Zdania w projekcie Tatoeba skupiają się na tłumaczeniu sposobu wyrażenia danej konstrukcji w innym języku, a niekoniecznie obsługiwanie **wszelkich możliwych** ekwiwalentów sytuacyjnych. Proponujemy - dla porządku i na prawach szerszej konwencji - przyjęcie rodzaju męskiego i liczby pojedynczej. Co do poziomu honoryfikatywnego (a więc zdania na "pan" vs. na "ty"), to język angielski nie pomoże, ponieważ zwykle nie stosuje takich rozróżnień. Jeśli możesz, tłumacz takie wątpliwe zdania z innych języków, które

rozdzielenie stylu mowy stosuj (np. rosyjski, francuski, japoński).

- Pamiętaj o interpunkcji i spacji! W polskim nie dajemy spacji przed przecinkiem, kropką, znakiem zapytania, wykrzyknikiem itd. Myślnik natomiast musi mieć spację przed i po. Nie używaj też wielokrotnego znaku zapytania i wykrzyknika (??? i !!!) - raz naprawdę wystarczy.

No i co ważne, nie bój się zwracać uwagi innym na popełnione błędy - w takiej liczbie zdań błędy są nieuchronne! Nie obrażaj się też, jeśli ktoś zwróci uwagę tobie!