Krzysztof Sitkowski

Kompresja tekstowa jako realizacja założeń ekonomii językowej

Językoznawstwo : współczesne badania, problemy i analizy językoznawcze 5, 153-158

2011

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

MUZEUM HISTORII POLSKI

Krzysztof Sitkowski

Kompresja tekstowa jako realizacja założeń ekonomii językowej

Ekonomię można określić jako umiejętność racjonalnego użytkowania czegoś. Umiejętne posługiwanie się słowem posiada wielkie znaczenie w procesie komunikacji, zwłaszcza jeżeli chodzi o racjonalne wykorzystywanie odpowiednich ilości środków wyrazu. Ekonomia językowa jest istotna ze względu na konieczność przekazywania coraz większych porcji informacji. Obecne tempo rozwoju ludzkości dostarcza coraz to nowych wiadomości. Niestety, dostępna ilość czasu nie uległa zmianie. Brak czasu sprawia, iż nadawca, aby w pełni wyrazić swoje myśli w sposób zrozumiały, musi odpowiednio ograniczyć ilość środków wyrazu. Najłatwiej uczynić to poprzez redukcję jednostek zbędnych, np. takich które odnoszą się do informacji już znanych odbiorcy.

Ekonomia językowa objawiać się może zwięzłością wypowiedzi. Według René Dubuca[1] zwięzłość jest właściwością języka. Intuicyjnie, użytkownik języka stara się wyrazić jak największą porcję informacji, przy zastosowaniu jak najmniejszej ilości środków językowych. Dzięki temu, w językach specjalistycznych, termin, który w swoim założeniu jest zwięzły, posiada przewagę nad pozostałymi jednostkami języka.

Odpowiedni poziom ekonomii językowej osiągnąć można przez stosowanie zwięzłości wypowiedzi, polegającej na niewystępowaniu w niej elementów zbędnych ze względu na jej funkcję komunikatywną i nie pełniących żadnych funkcji stylistycznych. Zwięzłość nieszkodząca zrozumiałości tekstu stanowi jego zaletę, ponieważ oszczędza nadawcy i odbiorcy wypowiedzi wysiłku oraz czasu potrzebnego do porozumienia się.

Do stosowania zasad ekonomii językowej nawoływali badacze, tacy jak Herbert Paul Grice czy Robin Lakoff. H. P. Grice w swoich maksymach komunikacyjnych postulował, aby wkład do konwersacji był na tyle informatywny, na ile jest to wymagane z punktu

[1] R. Dubuc, *Manuel pratique de terminologie*, Quatrième édition, Montréal 2002, s. 130.

widzenia aktualnych celów rozmowy, a także aby powstrzymać się od komunikowania większych ilości informacji, niż jest to wymagane. Dodatkowo, w maksymie relewancji nawoływał do mówienia ściśle na temat, a w maksymie sposobu – do unikania niepotrzebnej rozwlekłości[2].

Według R. Lakoffa[3] należy wyrażać się jasno, aby przekazywać informacje w możliwie szybki i efektywny sposób. Przestrzeganie powyższych zasad ma zapewnić skuteczność komunikacji.

Po omówieniu ekonomii językowej pora przejść do kolejnego konceptu – kompresji. Kompresję można scharakteryzować jako działanie, które ma na celu redukcję nadmiaru językowych środków wyrazu, tj. redukcję redundancji lub zmniejszenie objętości zastosowanych jednostek informacji przy jednoczesnym zachowaniu wartości informacyjnej.

Kompresja określona jest mianem tekstowej, ponieważ występuje w tekstach. Te natomiast są produktem języka, naturalnego lub specjalistycznego. Każdy konkretny tekst został wytworzony przez konkretną osobę na podstawie jej języka, dla wyrażenia konkretnej porcji wiedzy, która stanowi wynik analizy, obserwacji czy też innej czynności poznawczej[4].

Wyróżnia się trzy podstawowe typy kompresji tekstowej[5].

Kompresja logiczna polega na ograniczeniu objętości tekstu poprzez implicytne wyrażanie treści, tj. redukcję redundancji. Głównym założeniem kompresji logicznej jest występowanie w tekstach, zarówno ogólnych jak i specjalistycznych redundancji, czyli nadmiaru językowej organizacji wypowiedzi, przejawiającego się w występowaniu takich jej elementów, które nie służą bezpośrednio przekazywaniu informacji i nie są z punktu widzenia tej informacji konieczne. Redundancja prowadząca do straty czasu, energii, materiału i przestrzeni nazywana jest redundancją pustą. Natomiast redundancja, która powoduje rozproszenie uwagi, demotywację i zniechęcenie poprzez powtarzanie informacji zbędnych, nazywana jest redundancją zakłócającą[6]. Nadawca wiedząc, że pewne elementy jego komunikatu są znane odbiorcy, będzie mógł przekazać je implicytnie. W tym przypadku, konieczna jest pewność dotycząca odpowiedniego poziomu wiedzy językowej odbiorcy. Stosowana w komunikacji profesjonalnej, pomiędzy użytkownikami tego samego kodu, będącymi w stanie uzupełnić pominięte treści, ta forma kompresji może przyczynić się do znacznego przyspieszenia i usprawnienia komunikacji.

---

2 H. P. Grice, *Logic and conversation*, [w:] P. Cole, J. L. Morgan (red.), *Syntax and Semantics*, vol. 3. Speech Acts. New York, San Francisco, London 1975, s. 41–58.

3 R. Lakoff, *What you can do with words: Politeness*, *pragmatics*, *and performatives*, [w:] A. Rogers, B. Wall, J. P. Murphy (red.), *Procedings of the Texas Conference on Performatives*, *Presuppositions*, *and Implicatures*, Arlington 1977, s. 79–105.

4 S. Grucza, *„Teksty specjalistyczne": Językowe eksponenty wiedzy specjalistycznej*, [w:] Ł. Karpiński (red.), *Kulturowy i leksykograficzny obraz języków specjalistycznych*, „Języki Specjalistyczne", t. 8, Warszawa 2008, s. 181–193.

5 M. Górnicz, *Sposoby i granice kompresji terminów (na podstawie terminologii medycznej – nazwy nowotworów)*, „Poradnik Językowy", nr 2 (541), 1997, s. 25–32.

6 J. Zydroń, *Redundancja językowa w tekstach popularnonaukowych i materiałach glottodydaktycznych*, „Przegląd Glottodydaktyczny", nr 12, 1993, s. 25–35.

Ludzki mózg jest w stanie uzupełnić brakujące informacje na podstawie posiadanej wiedzy. Dzieje się to na tej samej zasadzie, co czytanie tekstu, z którego usunięto niektóre litery lub takiego, który jest niewyraźnie napisany.

Kompresja syntagmatyczna – zastosowanie w tekście odpowiednich środków składniowych, takich jak: zastąpienie zdań podrzędnych wyrażeniami przyimkowymi, stosowanie przyimków złożonych, rzeczowników dewerbalnych oraz utartych połączeń leksykalnych i syntaktycznych w celu zmniejszenia jego objętości. Za taką formę kompresji przyjąć można redukcję pleonazmów, tj. wyrazów treściowo pustych (m.in. *teren*, np. *na terenie Warszawy zbudowano nowy wieżowiec = w Warszawie zbudowano…*; *fakt*, np. *trzeba się liczyć z faktem ujemnego wpływu = ...z ujemnym wpływem*; *proces*, np. *metoda ważna w procesie nauczania = ...w nauczaniu*) oraz tzw. wyrażeń opisowych (np. *ulec zepsuciu = zepsuć się*).

Redukcja pleonazmów ma szczególne znaczenie w przypadku form zapożyczonych z innych języków, w których nagminnie pojawiają się niepotrzebne powtórzenia, np. wirus HIV – pełna forma zawiera już cząstkę *virus*. Tak samo jest w przypadku numeru PIN – *personal identification number*, pamięć RAM – *random access memory.*

Kompresja terminologiczna – zestaw operacji przeprowadzanych na terminach w celu zmniejszenia ich wielkości materialnej. Głównymi rodzajami kompresji terminologicznej są:

– kompresja skrótowcowa, polegająca na zastosowaniu form skróconych zamiast pełnych realizacji nazw i terminów. Zaletą tejże formy jest uproszczenie nazw trudnych i skomplikowanych;
– kompresja słowotwórcza, która zastępuje łańcuchy terminologiczne wyrazami złożonymi. Proces ten opiera się na łączeniu początkowych i końcowych segmentów wyrazów, wchodzących w skład wielowyrazowego pojęcia. Doskonałym przykładem zastosowania tego rodzaju kompresji jest teleskopowa jednostka terminologiczna. Z tą formą kompresji powiązany jest proces uniwerbizacji, czyli zastąpienie nazwy złożonej, najczęściej rzeczownika i opisującego go przymiotnika, wyrazem pojedynczym; *(samochód ciężarowy – ciężarówka*, *oddział porodowy – porodówka*, *port gazowy – gazoport)*;
– kompresja semantyczna, polegającą na usunięciu z terminu niektórych składników pojęcia. Zjawisko to możliwe jest dzięki istnieniu w terminie elementu, który pozwala na odtworzenie znaczenia składnika usuniętego; *(indukcyjny piec elektryczny*, *łukowy piec elektryczny* – cząsteczkę „elektryczny" można pominąć, ponieważ indukcja i łuk są zjawiskami wykorzystującymi właściwości prądu elektrycznego;
– kompresja morfologiczna bazującą na wykorzystaniu fałszywych morfemów powstałych na skutek nieprawidłowego dzielenia słów;
– kompresja semiotyczna, zakładająca zastosowanie symboli. Szczególnego znaczenia nabierają sztuczne kody semiotyczne[7].

[7] V. M. Lejczyk, *Optimalnaja dlina i optimalnaja struktura termina*, „Voprosy Jazykoznanija", nr 2, 1981, s. 63–72.

Do kompresji terminologicznej zaliczyć można także uniwerbizację i nominalizację. Zauważyć należy, iż samo użycie terminów określić można mianem kompresji. Termin jest skondensowaną formą informacji, zastępującą i będącą równą długiej definicji pojęcia. Według wytycznych polityki terminologicznej, opracowanych przez UNESCO[8], jedną z obecnie obowiązujących zasad tworzenia terminów jest zwięzłość, czyli ekonomia środków przekazu językowego. Również Juan Sager[9] stwierdził, iż *terminy powinny być zwięzłe i nie zawierać zbędnych informacji*.

Zastosowanie metod kompresji językowej posiada wiele zalet, między innymi:

– uproszczenie komunikatu;
– redukcja redundancji – nadawca może pozbyć się nieistotnych w jego przekonaniu elementów informacji. Pozwala to także na wyeliminowanie czynników, które mogłyby rozpraszać uwagę odbiorców;
– redukcja materialnej wielkości porcji informacji przy jednoczesnym zachowaniu wszelkich właściwości informacyjnych;
– uproszczenie i przyspieszenie notowania – ze względu na mniejszą liczbę użytych jednostek języka;
– zwiększenie możliwości zapamiętywania. Ta sama porcja informacji wyrażona jest za pomocą mniejszej liczby znaków. Znaki te są ze sobą ściśle powiązane;
– uproszczenie długich nazw i terminów – najlepszym przykładem jest język chemii – tri-nitro-toluen to trotyl lub TNT;
– wymuszanie funkcjonowania zasobów pamięciowych. Odbiorca, aby w pełni zrozumieć daną wiadomość, musi wykorzystać swoje zasoby pamięciowe, a także zdolność kojarzenia faktów.

Po zastosowaniu metod kompresji wartość informacyjna nie ulega zmianie. Komunikat zostaje pozbawiony informacji zbędnych, znanych odbiorcy lub takich, których brak nie będzie powodował destrukcji tekstu. Zakłada się, że odbiorca będzie w stanie uzupełnić luki powstałe po zastosowaniu kompresji, na podstawie posiadanej wiedzy. Podobny proces zachodzi podczas czytania tekstu, w którym opuszczono niektóre samogłoski.

Stopień skompresowania tekstu zależy oczywiście od poziomu wiedzy językowej odbiorcy. W komunikacji pomiędzy specjalistami, zastosowanie zabiegów kompresyjnych będzie o wiele większe. W komunikacji pomiędzy wykładowcą, a studentem, stopień ten będzie mocno ograniczony niższym poziomem wiedzy językowej tego drugiego. Do niestosowania pełni możliwości kompresji obligować także będzie rodzaj dyskursu. Zadaniem wykładowcy jest pomoc w zbudowaniu pewnej wizji konkretnego konceptu. Jednakże, nawet w takiej sytuacji, możliwe jest zastosowanie kompresji, np. w stosunku do porcji informacji już znanych studentowi.

O popularności zjawiska kompresji zarówno w językach specjalistycznych, jak i naturalnych świadczyć może rosnąca liczba słowników akronimów wszelkich typów np. witryna *Acronym Finder*[10], która zawiera definicje ponad czterech milionów jednostek skróconych.

---

8 UNESCO, *Guidelines for Terminology Policies. Formulating and implementing terminology policy in language communities*, Paris 2005.

9 J. C. Sager, *A practical course in terminology processing*, Amsterdam, Philadelphia 1990, s. 89.

10 Acronym Finder, http://www.acronymfinder.com/about.asp.

W slangu internetowym, który uznać można za skompresowaną postać języka, oferowane są nawet dzieła literackie. Jeden z brytyjskich operatorów sieci komórkowej zaoferował swoim klientom skrócone wersje dzieł literackich. Odbiorca otrzymuje SMS-a zawierającego skompresowaną wersję powieści. W procesie tworzenia owych streszczeń udział brali językoznawcy i literaturoznawcy, z prof. Johnem Sutherlandem na czele. Usługa *Text book service* obejmuje dzieła Williama Shakespeara, Johna Miltona, Charlesa Dickensa, Emily Brontë oraz wielu innych. Oczywiście, rozwiązania takiego nie można porównać z czytaniem całego dzieła i rozkoszowaniem się pełnią form literackich oraz kunsztem autora. Jednakże, skompresowane wersje dzieł literackich stanowić mogą doskonały bodziec do ćwiczenia pamięci.

Po co nam ekonomia językowa? Odpowiedzi na to pytanie może być kilka. Na pewno ekonomia językowa wpływa na przyspieszenie i usprawnienie komunikacji. Dzięki zastosowaniu odpowiednich procedur zaoszczędzić można czas, energię potrzebną do wyrażenia, przekazania i odbioru porcji informacji oraz materiały używane w trakcie tego procesu. Ekonomia językowa jest też na pewno efektem ewolucji języka. Z każdą chwilą rośnie ilość informacji, które powinny zostać przekazane, a czas na to przeznaczony nie zmienia się. Taka sytuacja zmusza nadawcę do wartościowania elementów komunikatu i wydzielania porcji informacji, które mogą zostać pominięte. Kolejnym powodem stosowania oszczędności środków wyrazu może być zwykłe lenistwo lub chęć podążania za najnowszymi trendami. Znaczący wpływ na liczbę metod i środków oszczędności ma rozwój komunikacji wspieranej komputerowo. Wystarczy powiedzieć, że SMS, którego długość wynosi 160 znaków, jest w pełni wystarczającym medium komunikacji. Jednakże, aby za pomocą 160 znaków przekazać rozbudowaną myśl, należy odpowiednio skondensować jej formę. Sztuka ta została opanowana przez użytkowników komunikacji wspieranej komputerowo (*Computer assisted Communication*), obfitującej w formy takie jak: kombinacje piktogramów, logogramów, form skróconych, opuszczeń liter, a zwłaszcza samogłosek, zapisu fonetycznego oraz jednostek zapożyczonych z języków obcych. Taki rodzaj komunikacji zyskał wielką popularność, a jego formy ewoluują w stronę ciągłego zmniejszania objętości komunikatu.

Ekonomia językowa oraz metody kompresji mogą znaleźć zastosowanie w procesie edukacji akademickiej. Konieczne byłoby wcześniejsze zbadanie poziomu wiedzy studentów, a także ciągłe monitorowanie zrozumienia konceptu. Z drugiej strony, sama kompresja może być wykorzystywana przy sprawdzaniu poziomu zrozumienia komunikatu. Studenci mieliby opisać poznany element wiedzy w postaci skompresowanej, a potem omówić go.

Kompresja może także znaleźć zastosowanie w przekładzie. Pośrednik językowy może pomijać pewne treści, które nie są istotne dla wartości informacyjnej danego komunikatu. Co więcej, przy osiągnięciu odpowiedniego poziomu znajomości z odbiorcą, tzw. zbieżności kodów, tłumacz może przedstawiać swoje tłumaczenie w postaci skompresowanej. Odbiorca, bazując na znajomości kodu, będzie w stanie rozszyfrować wiadomość.

Bibliografia

Acronym Finder, http://www.acronymfinder.com/about.asp.

Dubuc R., *Manuel pratique de terminologie*, Quatrième édition, Linguatech, Montreal 2002.

Górnicz M., *Sposoby i granice kompresji terminów (na podstawie terminologii medycznej – nazwy nowotworów)*, „Poradnik Językowy", nr 2 (541), 1997.

Grice H. P., *Logic and conversation*, [w:] Cole P., Morgan J. L. (red.), *Syntax and Semantics*, vol. 3. *Speech Acts*, Academic Press, New York, San Francisco, London 1975.

Grucza S., *„Teksty specjalistyczne": Językowe eksponenty wiedzy specjalistycznej*, [w:] Karpiński Ł. (red.), *Kulturowy i leksykograficzny obraz języków specjalistycznych*, „Języki Specjalistyczne", t. 8, Katedra Języków Specjalistycznych, Warszawa 2008.

Lakoff R., *What you can do with words: Politeness*, *pragmatics*, *and performatives*, [w:] Rogers A., Wall B., Murphy J. P. (red.), *Procedings of the Texas Conference on Performatives*, *Presuppositions*, *and Implicatures*, Center for Applied Linguistics, Arlington 1977.

Lejczyk V. M., *Optimalnaja dlina i optimalnaja struktura termina*, „Voprosy Jazykoznanija", nr 2, 1981.

Sager J. C., *A practical course in terminology processing*, Benjamins, Amsterdam, Philadelphia 1990, s. 89.

UNESCO, *Guidelines for Terminology Policies. Formulating and implementing terminology policy in language communities*, Infoterm, Paris 2005.

Zydroń J., *Redundancja językowa w tekstach popularnonaukowych i materiałach glottodydaktycznych*, „Przegląd Glottodydaktyczny", nr 12, 1993.