

JĘZYKOZNAWSTWO KRYMINALISTYCZNE W KRAJACH ANGLOJĘZYCZNYCH

Forensic linguistics in English-speaking countries

Językoznawstwo kryminalistyczne (inaczej: lingwistyka kryminalistyczna) zostało wyodrębnione stosunkowo niedawno jako część lingwistyki stosowanej¹. W najprostszym ujęciu definiuje się je jako dziedzinę wykorzystywaną do celów wykonywania czynności śledczych i prawnych przy zastosowaniu szeroko pojętej wiedzy językowej² (i korzystaniu z dorobku takich gałęzi językoznawstwa jak psycholingwistyka, socjolingwistyka, dialektologia, leksykografia), w przypadku gdy w sprawie pojawia się jakikolwiek twór językowy, bez względu na jego formę lub treść (zatem może badać teksty pisane i mówione, pisane odręcznie, drukowane, dostępne w formie elektronicznej).

W literaturze przedmiotu podaje się, iż po raz pierwszy pojęcie to (w j. ang. *forensic linguistics*) zostało użyte w 1968 roku przez szwedzkiego profesora językoznawstwa pracującego na University of London – Jana Svartvika³ – w publikacji *The Evans Statements: A Case For Forensic Linguistics*⁴, w której Svartvik poddał analizie zeznania niejakiego Timothy’ego Johna Evansa uznanego za winnego zabójstwa żony oraz córki, a pośmiertnie ułaskawionego dzięki wynikom analizy językowej złożonych przez niego zeznań. Zainteresowanie tematem językoznawstwa kryminalistycznego rośnie w ostatnich latach, głównie ze względu na wzrost liczby przestępstw popełnianych z użyciem lub za pośrednictwem języka⁵ (takich jak np. wysyłanie anonimów, fałszerstwa te-

¹ J. Stawnicka, *Językoznawstwo w służbie wymiaru sprawiedliwości*, „Kwartalnik Policyjny” 2015, nr 2, <http://kwartalnik.csp.edu.pl/kp/archiwum-1/2015/nr-22015/2681>, *Jezykoznawstwo-w-sluzbie-wymiaru-sprawiedliwosci.html*, dostęp 29.06.2017.

² Ibidem.

³ Uznanego za „ojca” lingwistyki kryminalistycznej.

⁴ J. Olsson, *Word Crime. Solving Crime through Forensic Linguistics*, Bloomsbury Academic, London 2009, s. 1.

⁵ Magda Olszewska w artykule *Anonim jako ślad kryminalistyczny* („Prokuratura i Prawo” 2014, nr 11–12, s. 130) zauważa, iż „na ogólną ilość 178 przestępstw podstawowych aż 51 może być popełnione wyłącznie z pomocą języka lub pisma”.

stamentów, umów, plagiat, stalking) oraz przede wszystkim wzrost znaczenia komunikacji internetowej (np. cyberstalking, ustalanie autorów wpisów na forach internetowych, ustalanie znaczenia słów użytych na czatach internetowych). W związku z powyższym coraz większym zainteresowaniem cieszy się także dział lingwistyki kryminalistycznej nazywany profilowaniem socjolingwistycznym, którego zadaniem jest przede wszystkim próba identyfikacji nieznanego autora badanego tekstu przez określenie jego pochodzenia geograficznego, płci, wieku, wykształcenia⁶. Jednak pomimo wzrostu liczebności spraw, w których zastosowanie metod badawczych językoznawstwa kryminalistycznego może być pomocne, dziedzina ta jest wciąż jeszcze w pewien sposób niedoceniana w polskim wymiarze sprawiedliwości. Wśród przedstawicieli organów ścigania często można spotkać się nawet z nieświadomością jej istnienia, a analizy językowe traktuje się zazwyczaj jedynie jako pomocniczy fragment klasycznej ekspertyzy pismoznawczej.

Natomiast zupełnie inaczej rozwój tej dziedziny przebiega w Wielkiej Brytanii i Stanach Zjednoczonych. Językoznawstwo kryminalistyczne w tych krajach funkcjonuje zdecydowanie najdłużej i równocześnie najbardziej się rozwija, z pewnością dzięki licznym sprawom, w których je z sukcesem stosowano (jak wspomniana sprawa Evansa w Wielkiej Brytanii lub też analiza językowa w sprawie Unabombera w Stanach Zjednoczonych⁷). Lingwistyka kryminalistyczna po dziś dzień cieszy się tam dużym poparciem oraz uznaniem i jest traktowana w zasadzie jak każda inna dziedzina kryminalistyki, a zagadnienia z nią związane są popularyzowane w książkach i serialach. Warto tu przytoczyć choćby popularny w Stanach Zjednoczonych serial *Criminal Minds* (w Polsce – *Zabójcze umysły*), opowiadający o pracy wydziału analiz behawioralnych FBI. W ósmym i dziewiątym sezonie serialu wprowadzono do fabuły postać językoznawcy kryminalistycznego pomagającego agentom rozwiązywać liczne sprawy kryminalne z pomocą analizy lingwistycznej. Natomiast spośród książek trzeba wspomnieć o publikacji Johna Olssona pt. *Word Crime: Solving Crime through Forensic Linguistics* (Bloomsbury Academic, London, 2009), w której autor w przystępny sposób, prostym językiem opisuje kilkanaście różnorodnych spraw ze swojej praktyki, metody badawcze, którymi się w nich posłużył, oraz podstawowe zagadnienia i problemy związane ze stosowaniem lingwistyki kryminalistycznej. Taka popularyzacja sprawia, iż coraz więcej osób jest zainteresowanych tematyką językoznawstwa kryminalistycznego, co z ko-

⁶ J. Stawnicka, *Biegły językoznawca, czy lingwista kryminalistyczny*, 2013 <http://www.popravny.pl/index.php?itemid=5167&catid=36>, dostęp 29.06.2017.

⁷ Więcej na temat analizy językowej w tej sprawie: <http://www.pbs.org/speak/seatosea/americanvarieties/DARE/profiling/>, dostęp 29.06.2017.

lei przekłada się na przywiązywanie dużej wagi do wykształcenia nowych pokoleń językoznawców kryminalistycznych, przede wszystkim dzięki licznie powstającym ogólnodostępnym kursom internetowym z tego zakresu, czy też coraz bardziej popularnemu, specjalnemu kierunkowi studiów – *forensic linguistics*⁸. Dodatkowo powstają liczne stowarzyszenia, których zadaniem jest zrzeszanie ekspertów z tej dziedziny, ujednocianie i wprowadzanie nowych metod badawczych oraz opracowywanie kodeksów etyki (np. International Association of Forensic Linguists, Forensic Linguistics Institute). Przy tym warto zaznaczyć, iż rozumienie pojęcia *forensic linguistics* w Wielkiej Brytanii i Stanach Zjednoczonych jest znacznie szersze niż na gruncie polskim, a co za tym idzie, obejmuje znacznie więcej zastosowań. Dosyć uproszczone, ale i w pełni uzasadnione stwierdzenie o lingwistyce kryminalistycznej na tamtych terenach głosi, iż lingwistyka kryminalistyczna zajmuje się badaniem wszelkich związków języka z prawem, łączących się z zastosowaniem metod językoznawczych do analizy zagadnień prawnych⁹. Wobec tego opisując tę dziedzinę, często mówi się po prostu o *language and the law*. I zgodnie z tą tendencją, w większości książek, które mają w tytule *forensic linguistics*, można się jeszcze zapoznać, oprócz zagadnienia śladu językowego jako dowodu, z wieloma innymi zastosowaniami. Zastosowania te można podzielić na trzy szersze kategorie:

- 1) analizę języka prawa pisanego;
- 2) analizę języka procesu sądowego;
- 3) wykorzystanie śladu językowego jako dowodu¹⁰.

Należy zaznaczyć, iż w Polsce dwie pierwsze kategorie (język prawa pisanego oraz język procesu sądowego) włącza się do zakresu juryslingwistyki, w obszarze zainteresowania językoznawstwa kryminalistycznego zaś pozostawia się jedynie analizy śladów językowych w charakterze dowodów w sprawie.

Skupiając się na szerokim rozumieniu pojęcia *forensic linguistics*, w każdej z powyżej przedstawionych kategorii można wyodrębnić jeszcze liczne, skonkretyzowane zastosowania. I tak w obrębie badań języka prawa pisanego analizuje się cechy, jakimi powinno charakteryzować się słownictwo wszelkich aktów prawnych i pism prawniczych, by były one powszechnie zrozumiałe, oraz poszukuje się metod interpretacji oraz ulepszenia języka aktów prawnych,

⁸ Przykładowe oferty kursów i kierunków studiów z zakresu językoznawstwa kryminalistycznego: <http://www.thetext.co.uk/courses.html>, <http://www.aston.ac.uk/study/postgraduate/taught-programmes/languages-social-sciences/forensic-linguistics-ma/>, dostęp 29.06.2017.

⁹ J. Olsson, *Forensic Linguistics: Second Edition*, Continuum International Publishing Group, London–New York 2008, s. 3.

¹⁰ *Language of the written law, language use in forensic and judicial processes, linguistics evidence*.

aby charakteryzował się przejrzystością, precyzyjnością, jednoznacznością, spójnością¹¹. Te same postulaty próbuje się włączyć do konstrukcji wszelkich umów, opartych na języku prawnym, a przeznaczonych dla zwykłych obywateli (w tym do treści ostrzeżeń umieszczanych na produktach konsumenckich). Ważnym elementem tego działu jest także badanie problematyki dokonywania tłumaczeń wszelkiego rodzaju tekstów prawnych na inne języki, tak aby zachowując różnice językowe i przede wszystkim odmienność w zakresie przyjętych systemów prawnych, możliwie najpełniej odwzorować rzeczywistą treść aktu¹². Natomiast w zakresie dyskursu na sali sądowej, a w tym analiz wszelkich aktów komunikacji mniej lub bardziej związanych ze stosowaniem prawa, szczególne zainteresowanie budzą zagadnienia przesłuchań policyjnych (z naciskiem na ich taktykę oraz na problem fałszywego przyznania się do winy), interakcji mówca–słuchacz podczas przyjmowania zgłoszeń na policyjny telefon alarmowy¹³, komunikacji w relacji prawnik–klient oraz charakterystycznych cech języka używanego na sali sądowej¹⁴, głównie w kontekście przesłuchań świadków (z największym naciskiem na taktykę przesłuchań szczególnych), a także specyfiki językowej mów końcowych. W tym zakresie szczególną wagę przykłada się do sposobów formułowania i zadawania pytań, zagadnienia przejawiania autorytetu przedstawicieli wymiaru sprawiedliwości w używanym języku oraz trudności w rozumieniu języka prawa przez zwykłych obywateli (*lay participants*). W zakresie policyjnych przesłuchań w Stanach Zjednoczonych duży nacisk kładzie się także na ocenę stopnia zrozumienia przez przesłuchiwanego tzw. praw Mirandy (które może być zakłócone przez takie czynniki jak obniżone zdolności poznawcze, bycie rodzimym użytkownikiem innego języka, młody wiek¹⁵), aby na tej podstawie móc ocenić autentyczność przyznania się do winy przez osobę przesłuchiwaną¹⁶. Istotnym

¹¹ „Clarity, precision, unambiguity and inclusiveness”, w: M. Coulthard, A. Johnson, *The Routledge Handbook of Forensic Linguistics*, Routledge, New York 2013, s. 38.

¹² Często tłumaczenie bezpośrednie jest niemożliwe ze względu na to, iż z pozoru identyczne słowa mogą mieć w różnych językach inne konotacje i tłumaczenia. Za: M. Coulthard, A. Johnson, *The Routledge Handbook...*, op. cit., s. 91.

¹³ Szczególnie ciekawą kwestią jest analiza tzw. *emergency calls* pod kątem oceny powagi zgłaszanego zdarzenia oraz możliwości ewentualnego udziału w zdarzeniu (sprawstwa) na podstawie słownictwa użytego przez zgłaszającego. Za: M. Coulthard, A. Johnson, *The Routledge Handbook...*, op. cit., s. 95–110.

¹⁴ Ibidem, s. 199–230.

¹⁵ W szczególności badanie kompetencji językowych nastolatków w kontekście zrozumienia przysługujących im praw przedstawionych przed przesłuchaniem. Więcej na ten temat w: P. Tiersma, L. Solan, *The Oxford Handbook of Language and Law*, Oxford University Press, New York 2012, s. 299–311.

¹⁶ J. Olsson, *Forensic Linguistics...*, op. cit., s. 5–6.

obszarem jest także badanie specyfiki subjęzyków¹⁷ policjantów, sędziów, prokuratorów, adwokatów, więźniów, przestępców i innych grup w różny sposób związanych z prawem pod kątem specyfiki używanego słownictwa, struktur gramatycznych, struktury zdania, stylistyki celem identyfikacji przedstawicieli tych grup na podstawie wytworzonego przez nich tekstu. Coulthard¹⁸, poddawszy analizie kilka przykładów rejestru policyjnego, zauważa wiele podobieństw w sposobie używania języka przez wszystkich funkcjonariuszy policji. W języku angielskim bardzo charakterystyczną cechą takich zeznań jest między innymi specyficzne umiejscowienie łącznika *then* w strukturze zdania, który w mowie potocznej zwykle występuje przed podmiotem zdania (np. *Then I...*), natomiast w przypadku rejestru policyjnego zawsze będzie umiejscowiony po podmiocie (np. *I then...*). Co więcej, trudno takie użycie w sposób konsekwentny imitować¹⁹. Zatem po zidentyfikowaniu wielu przypadków występowania takiej struktury w jakiejś wypowiedzi, nieprzeplatanych dodatkowo potocznym użyciem, z dużą dozą prawdopodobieństwa można stwierdzić, iż autorem badanego tekstu jest policjant²⁰. Analizując język protokołów przesłuchań, można dodatkowo dokonywać oceny stopnia zmęczenia i zestresowania u osoby przesłuchiwanej. Dowodem na występowanie takiej „emocjonalnej niestabilności” przesłuchiwanego ma być na przykład zwiększona proporcja czasowników do przymiotników w protokole przesłuchania w porównaniu z innymi wypowiedziami tej samej osoby²¹. Równie częstym problemem spotykanym podczas oceny wiarygodności protokołów są też lingwistyczne oznaki dokładnego przepisywania ich treści z innych wypowiedzi podejrzanego. Takie analizy języka protokołów przesłuchań w wielu sprawach przyczyniły się do ułaskawienia oskarżonych, którzy rzekomo przyznali się w protokole do winy, a badania lingwistyczne następnie ujawniły dokonywanie manipulacji przy treści ich oświadczeń²².

¹⁷ Subjęzyk – „jedna z odmian języka ogólnego, którą posługuje się grupa ludzi wydzielona z ogółu na podstawie takiego kryterium jak np. pochodzenie terytorialne, płeć, wiek, wykształcenie, zawód, stan zdrowia psychicznego”. Cytat za: <http://prawouam-stp.home.amu.edu.pl/s.htm>, dostęp 29.06.2017.

¹⁸ M. Coulthard, A. Johnson, *An Introduction to Forensic Linguistics. Language in Evidence*, Routledge, Abingdon–New York 2007, s. 76–80.

¹⁹ Jak wykazał Coulthard w swoich badaniach, przeciętny człowiek, próbując naśladować tę cechę, zawsze będzie nieświadomie wtrącał do wypowiedzi także potoczną strukturę zdania z *then*, policjanci zaś konsekwentnie trzymają się tej specyficznej. Więcej w: *ibidem*, s. 77.

²⁰ Co może być także przydatne przy profilowaniu autorów anonimów, zob. sprawa męczyzny zatruwającego puszki z jedzeniem, w: M. Olszewska, *op. cit.*, s. 135–137.

²¹ „Jeśli iloraz czasownik/przymiotnik uzyskany z protokołu przyznania się był znacznie większy niż w innych wypowiedziach oskarżonego, to można wyciągnąć wniosek, że przyznanie się nastąpiło w stanie znacznej emocjonalnej niestabilności”, w: B. Hołyst, *Psychologia kryminalistyczna*, LexisNexis, Warszawa 2009, wydanie 3, r. LXXVII, s. 1327.

²² Tak też w sprawie T. Evansa (1949) oraz Dereka Bentleya (1952).

Przechodząc do trzeciej kategorii, tj. zagadnienia śladu językowego jako dowodu, wśród najważniejszych zastosowań językoznawstwa kryminalistycznego wymienia się:

- przypisanie lub porównanie autorstwa (anonimów, SMS-ów, wiadomości e-mail itp., w tym sprawy o przywłaszczenie autorstwa),
- ustalenie intencji nadawcy komunikatu (przede wszystkim w sprawach o groźby karalne, naruszenie dóbr osobistych),
- ustalenie znaczenia użytych w tekście słów (analizy wyrażen slangowych grup przestępczych, narkotykowych, terrorystycznych),
- profilowanie językowe autora tekstu (czyli zarysowanie charakterystyki socjolingwistycznej autora przez określenie jego wieku, płci, pochodzenia terytorialnego, wykonywanego zawodu, poziomu wykształcenia), szczególnie przydatne do typowania autorów anonimów pogroźkowych i znieważających, żądań okupu, listów pożegnalnych oraz identyfikacji stalkerów i cyberstalkerów.

W celu praktycznego wdrażania wyżej wymienionych zastosowań praktyka wypracowała szereg lingwistycznych metod badawczych pozwalających na realizowanie tych zadań. W tym zakresie wyróżnia się:

a) metody ilościowe (*quantitative methods*) – zazwyczaj wspomagane komputerowo, oparte głównie na statystyce i wyliczeniach matematycznych, polegające na analizie ilościowego występowania pewnych cech w obrębie badanej wypowiedzi lub też całej populacji. W literaturze angielskiej zwane również „mierzeniem stylu” (*the measurement of style*);

b) metody jakościowe (*qualitative methods*) – bardziej opisowe, korzystające z założeń nauk humanistycznych, a więc językoznawstwa ogólnego, psycholingwistyki, dialektologii i innych pokrewnych metod. Badające komunikat jako całość powstałą pod wpływem wielu nakładających się na siebie czynników. Oparte głównie na analizie cech subiektywnie wybranych przez eksperta; zwane też „opisem stylu” (*description of style*).

W analizach ilościowych najczęściej korzysta się z dorobku stylometrii i lingwistyki korpusowej. Stylometria jest metodą polegającą na statystycznej charakterystyce stylu autora, której dokonuje się głównie przez ustalenie liczbowych rozkładów stylemów (elementarnych jednostek językowo-stylistycznych)²³ w badanych tekstach. W najprostszym ujęciu analiza sprowadza się do liczbowego przedstawiania najczęstszych cech lingwistycznych występujących w tekście²⁴, w tym przede wszystkim:

²³ K. Polański, *Encyklopedia językoznawstwa ogólnego*, Zakład Narodowy im. Ossolińskich, Wrocław 1999, s. 568.

²⁴ B. Hołyst, op. cit., r. LXXVII, s. 1326.

- częstości występowania określonych słów i części mowy;
- bogactwa słownictwa (różnorodności używanych wyrazów – głównie synonimów);
- średniej długości słów (mierzonej przez zliczanie liczby liter lub sylab w obrębie wyrazu);
- proporcji użycia różnych części mowy w tekście (proporcji czasowników do rzeczowników, liczby przymiotników określających rzeczowniki itp.);
- średniej długości zdania (określanej przez liczbę słów – tokenów w zdaniu);
- sposobów organizowania danych w obrębie zdania²⁵.

Można wyróżnić dwa podejścia do analizy stylometrycznej. Pierwsze z nich skupia się na analizie samej struktury języka, a więc cech stylistycznych, słownictwa, gramatyki, składni, pisowni (jak we wspomnianej sprawie T. Evansa). Natomiast drugie, bliższe klasycznemu rozumieniu stylometrii, obejmuje czysto matematyczne pomiary częstotliwości słów i ich układu w tekście oraz wyliczanie proporcji poszczególnych wyrazów do wszystkich wyrazów w tekście, w tym tzw. *type-token ratio*, czyli proporcje typów słów występujących w tekście (słowa powtarzające się liczone są tylko raz) do ogólnej liczby wszystkich słów, także długość zdań i organizację danych w ich obrębie, liczbę sylab na 100 słów, liczbę przyimków, proporcję czasowników do rzeczowników, czasowników do przymiotników itd. Przy tym zwraca się uwagę na dużą wartość wyrażen synsemantycznych w tekstach, a więc takich, które pełne znaczenie otrzymują dopiero w połączeniu z inną częścią mowy, samodzielnie zaś nic nie znaczą, np. przyimki, spójniki, partykuły (w języku angielskim *function words* – np. and, a(n), the) – takie wyrazy pojawiają się w zasadzie w każdym tworze pisanim, a człowiek stosuje je w większości przypadków podświadomie, we względnie jednolity, właściwy sobie sposób, dlatego też są one wartościową cechą identyfikacyjną. Po zidentyfikowaniu i liczbowym zestawieniu wybranych cech językowych w najprostszej wersji badań stylometrycznych tworzy się dwie listy frekwencyjne cech – współwystępujących w tekście dowodowym i porównawczym oraz różniących oba teksty, a tym samym określa się istniejące między tekstami podobieństwa i różnice. Im zaś więcej podobieństw w tym zakresie zostanie ujawnionych, tym większe będzie prawdopodobieństwo, że oba teksty stworzyła jedna osoba.

²⁵ M. Zaško-Zielińska, *Listy poźegnalne. W poszukiwaniu lingwistycznych wyznaczników autentyczności tekstu*, Wydawnictwo Quaestio, Wrocław 2013, s. 20.

NH 2 (a): 44 sentences of 867 words=19.7 words/sentence
 NH 2 (b): 47 sentences of 789 words=16.8 words/sentence
 NH 2 (c): 18 sentences of 280 words=15.6 words/sentence

Ryc. 1. Przykład zastosowania analizy stylometrycznej w sprawie T. Evansa – oszacowanie liczby słów występujących w obrębie jednego zdania we fragmentach zeznania Evansa z dnia 2 grudnia 1949 r.

Źródło: J. Svartvik, *The Evans Statements. A case for forensic linguistics*, part 1, „Acta Universitatis Gothoburgensis” 1968, s. 20.

Niektóre metody stylometryczne skupiają się także na wyrażeniach autosemantycznych (ang. *lexical words*), a więc samoznających, niosących w sobie autonomiczne znaczenie. Tego rodzaju analizy przydają się na przykład w badaniach porównawczych tekstów pod kątem określenia bogactwa słownictwa ich autorów, a dzięki temu możliwości oceny, który z nich mógł w danym wypadku stworzyć kwestionowany tekst. W tym celu Winter i Woolls zaproponowali następujący algorytm:

$$100 \times \log N / (1 - LV_1 / V)$$

gdzie:

- N** – liczba wszystkich słów w tekście (tokenów)
- LV₁** – łączna liczba wyrazów autosemantycznych występujących tylko raz (tzw. *hapax legomena*)
- V** – łączna liczba poszczególnych typów słownictwa występujących w tekście.

Aby wykazać przydatność wyżej opisanej formuły, jej autorzy przeprowadzili analizę fragmentów kilku rozdziałów powieści, która została napisana wspólnie przez dwie osoby. W celu dokładniejszej weryfikacji metody do analizy dołączono jeszcze fragment dzieła napisanego tylko przez jedną z nich. W tabeli poniżej przedstawiono porównanie wskaźników różnorodności leksykalnej poszczególnych rozdziałów powieści oraz fragmentu dzieła pojedynczego autorstwa, podzielonego na trzy próbki kontrolne:

	<i>Av. sentence length</i>	<i>Lex. richness</i>
Ch 1	33.3	875
Ch 3	51.4	913
Ch 5	43.2	810
Ch 33	30.1	928
Ch 2	23.6	765
Ch 4	12.5	624
Ch 32	12.1	709
Ch 28	13.9	832
Ch 29	19.6	785
Ch 30	19.4	884
Ch 31	15.9	822
Ctrl 1	23.9	703
Ctrl 2	14.7	735
Ctrl 3	17.0	693

Ryc. 2. Przykład analizy tekstów metodą Wintera i Woollsa

Źródło: M. Coulthard, A. Johnson, *An Introduction to Forensic Linguistics...*, op. cit., s. 166.

Rozdziały wykazujące największą zbieżność bogactwa leksykalnego zestawiono w grupach. W efekcie oceniono, iż rozdziały nr 1, 3, 5, 33 prawdopodobnie należą do innego autora niż rozdziały o numerach 2, 4, 32 wraz z trzema próbkami kontrolnymi. Ponadto, jak można zauważyć, wyniki rozdziałów od nr. 28 do 31 zawierają się pomiędzy wartościami poprzednich dwóch grup – eksperci ocenili tutaj, iż taka sytuacja może być skutkiem wspólnego ich napisania przez dwóch autorów²⁶. Otrzymane wyniki okazały się w zdecydowanej większości prawdą.

W latach 90. XX wieku Morton i Michaelson przedstawili metodę ilościową opartą na statystyce, którą od złączenia angielskich słów *cummulative* i *sum* nazwano metodą CUSUM. Jej innowacyjność polegała na skupieniu badań przede wszystkim w granicach poszczególnych zdań tekstu, a nie na tekście jako całości, jak czyniła większość dotychczasowych metod. W obrębie każdego zdania zliczano częstotliwość pojawiania się rzeczowników, słów zaczynających się od samogłosek, słów składających się z trzech lub czterech liter oraz słów dwu- lub trzyliterowych²⁷. Następnie sumy tych poszczególnych elementów odnoszono do długości całego zdania, w którym występowały, określonej na podstawie liczby wszystkich wyrazów występujących w jego obrębie. Każdy z tych elementów był następnie odnoszony do długości zdania, w którego obrębie wystąpił, i dodatkowo do tekstu jako całości, a ostateczny wynik był sumą częstotliwości występowania danej cechy zarówno w obrębie jednego zdania, jak i całości tekstu. Aby lepiej wyjaśnić tę metodę, warto posłużyć się przykładem podanym przez Coultharda:

²⁶ M. Coulthard, A. Johnson, *An Introduction to Forensic Linguistics...*, op. cit., s. 166.

²⁷ Ibidem, s. 167.

Sentence	Total words	Cusum words	23lw	Cusum 23lw
1	20	+8	7	+2
2	12	+8	5	+2
3	6	+2	3	0

Ryc. 3. Przykład zastosowania metody sum skumulowanych (CUSUM)

Źródło: M. Coulthard, A. Johnson, *An Introduction to Forensic Linguistics...*, op. cit., s. 167.

Na rycinie 3 przedstawiono wyniki CUSUM dla trzech zdań tekstu, w których średnia liczba słów w zdaniu wynosiła 12, a słów dwu- lub trzyliterowych było średnio 5 na każde zdanie. Jak widać, trzy wybrane do analizy zdania składają się kolejno z 20, 12 i 6 słów. Na podstawie tych danych wyjściowych oblicza się następnie wartości CUSUM (w tym przypadku kolejno 8,8,2 dla wszystkich słów i 2,2,0 dla słów dwu- lub trzyliterowych). Po obliczeniu sum wyniki przedstawia się w formie wykresu. W przypadku pojedynczego autorstwa wykresy powinny nachodzić na siebie, natomiast w miejscach, które zostały stworzone przez więcej niż jedną osobę – być rozbieżne. Nie ma wątpliwości, iż w tej metodzie (jak też w większości metod stylometrycznych) głównym środkiem dowodzenia jest skomplikowana, czysta statystyka i nie potrzeba rozległej wiedzy językowej, aby móc przeprowadzić takie badania. Stąd też tego rodzaju metody badawcze, jako nieuwzględniające złożoności czynników towarzyszących powstaniu komunikatu (np. funkcji danej wypowiedzi, rodzaju kontaktu, tematu i miejsca rozmowy) i dorobku dyscyplin językoznawczych, są coraz częściej zastępowane analizami jakościowymi. Jednak pomimo tych problemów i wątpliwych wyników wskazań niektórych metod ilościowych obecnie funkcjonuje także wiele wiarygodnych podejść badawczych w tym zakresie, skupiających się na różnych cechach lingwistycznych i stosujących rozmaite algorytmy²⁸. Aby usprawnić i zautomatyzować analizy, tworzy się programy komputerowe specjalnie przeznaczone do przeprowadzania badań stylometrycznych, takie jak np. ALIAS²⁹ oferujący wiele modułów nakierowanych na skonkretyzowane zadania (ustalanie autorstwa, profilowanie językowe), z możliwością dokonania analiz w kilku językach. Co więcej, w Internecie można znaleźć wiele darmowych programów stylometrycznych, dzięki którym w domowym zaciszu, bez konieczności korzystania z pomocy profesjonalistów możemy dokonać prostej analizy ilościowej wybranego przez nas i następnie wprowadzonego na stronę internetową tekstu. Dostępny jest na przykład pro-

²⁸ Wyczerpujące omówienie kilku takich metod w: G. McMenamin, *Forensic Linguistics. Advances in Forensic Stylistics*, CRC Press, Boca Raton 2002, s. 151–175.

²⁹ <http://aliastechnology.com/alias-overview/>, dostęp 29.06.2017.

gram liczący wszystkie słowa we wprowadzonym tekście oraz oszacowujący automatycznie *type-token ratio*, który dodatkowo także konfiguruje listę wszystkich występujących w tekście słów wraz ze wskazaniem liczby wystąpień każdego z osobna w całym komunikacie³⁰. Można również znaleźć programy, których założeniem jest wskazanie płci autora badanego tekstu na podstawie frekwencji użytych w nim słów³¹. Należy jednak pamiętać, że możliwość przeprowadzenia analizy stylometrycznej jest obwarowana wieloma ograniczeniami, takimi jak przede wszystkim minimalna długość tekstów, które można poddać tym badaniom. Uważa się, iż miarodajne wyniki uzyskuje się tylko w przypadku tekstów liczących powyżej 1000 słów, co z góry dyskwalifikuje większość tekstów pozostających w zainteresowaniu wymiaru sprawiedliwości (anonimy, SMS-y)³².

Drugim rodzajem analiz ilościowych są analizy korpusowe, polegające na porównywaniu kwestionowanej wypowiedzi z normą zwyczaju językowego danej społeczności, która została ujęta w korpusach językowych. Korpus językowy jest zbiorem tekstów, w którym zawarte są typowe użycia słów i konstrukcje zdaniowe oraz informacje o ich znaczeniu i funkcji³³. Można wyróżnić korpusy ogólne (Narodowy Korpus Języka Polskiego, korpus PWN), korpusy specjalistyczne (Kartoteka Dokumentów Anonimowych, Polski Korpus Listów Pożegnalnych) i korpusy rozumiane jako zbiory tekstów porównawczych do konkretnej sprawy. Dzięki zaangażowaniu w analizę lingwistyczną takich narzędzi można przede wszystkim ustalić częstotliwość wybierania danej formy lub struktury językowej w społeczeństwie oraz najczęściej nadawane im znaczenie, następnie odnosząc te dane do struktury konkretnego śladu językowego zawartego w tekście dowodowym. Wychodzi się tu z założenia, iż im rzadziej w społeczeństwie występuje użycie badanego słowa lub struktury w takiej samej formie bądź kontekście, tym większą wartość identyfikacyjną mają one w analizie językowej. Analizując zaś ślad językowy odbiegający od przyjętej powszechnie normy, łatwiej zawęzić krąg jego potencjalnych autorów. W literaturze angielskiej taka oryginalność struktury związana jest z teorią *markedness*, w której cechy współwystępujące w wypowiedziach wielu autorów są określane jako *unmarked*, natomiast te wysoce niepowtarzalne, niestandardowe

³⁰ <http://nlpdotnet.com/Services/Fingerprinter.aspx>, dostęp 29.06.2017.

³¹ <http://www.hackerfactor.com/GenderGuesser.php>, dostęp 29.06.2017.

³² Jako uzasadnienie tego ograniczenia podnosi się fakt, że w krótszych tekstach o wiele trudniej ustalić występowanie „nawyków wystarczająco częstych dla uzyskania miarodajnych wyników metod statystycznych”, w: B. Hołyst, op. cit., r. LXXVII, s. 1328.

³³ Strona internetowa Narodowego Korpusu Języka Polskiego: <http://nkjp.pl/>, dostęp 29.06.2017.

na tle języka ogólnego, określa się jako *marked*³⁴. W zakresie badań korpusowych takimi niestandardowymi cechami mogą być na przykład błędy językowe i kolokacje, które według niektórych teorii są tworzone w sposób najbardziej podświadomy³⁵.

Interesująca wydaje się obecnie występująca tendencja do wykorzystywania w charakterze korpusów popularnych wyszukiwarek internetowych (np. Google). Sposób ich działania jest bardzo zbliżony do korpusów klasycznych – ekspert rozpoczyna analizę od wpisania badanej frazy w pole wyszukiwania, następnie odnotowuje liczbę trafień stron, na których badana fraza wystąpiła w takiej samej formie jak w tekście dowodowym, a ostatecznie na tej podstawie oceniana jest popularność użycia badanego zwrotu w obrębie społeczeństwa. W ten sam sposób można również wyszukiwać najczęściej wybierane kolokacje wyrazów. Rycina poniżej przedstawia wynik analizy językowej, której zadaniem było ustalenie autorstwa anonimów otrzymywanych przez osoby występujące w roli świadków w sprawie o zgwałcenie. Równocześnie świadkowie otrzymywali listy od mężczyzny, który został w tej sprawie oskarżony. Obie grupy tekstów wykazywały wiele podobieństw, jednak najbardziej charakterystycznym (zatem określanym jako *marked*) było opisywanie zdarzeń niezgodnie z ich porządkiem chronologicznym (*I am and was*), co, jak się okazało, nie jest powszechnie stosowaną w społeczeństwie strukturą. Aby to udowodnić, ekspert zestawiał w tabeli wyniki uzyskane podczas wpisywania do przeglądarki internetowej różnych kombinacji tego wyrażenia zarówno w poprawnym porządku chronologicznym, jak i odwróconym:

'I was and am'	155,000
'I am and was'	20,600
'he was and is'	174,000
'he is and was'	21,300
'she was and is'	55,100
'she is and was'	727
'we were and are'	50,300
'we are and were'	904
'they were and are'	122,000
'they are and were'	31,800

Ryc. 4. Wyniki analizy struktury zdaniowej uzyskane przy użyciu przeglądarki internetowej

Źródło: J. Olsson, *Word Crime...*, op. cit., s. 141.

³⁴ J. Olsson, *Forensic Linguistics...*, op. cit., s. 45–46.

³⁵ D. Wright, *Stylistics versus Statistics: A corpus linguistic approach to combining techniques in forensic authorship analysis using Enron emails*, The University of Leeds School of English, wrzesień 2014, s. 27–30.

Dzięki liczbie „trafień” można zauważyć, iż zdecydowanie bardziej powszechne jest używanie czasu przeszłego przed użyciem czasu teraźniejszego, użycie odwrotne zaś występuje w porównaniu niezwykle rzadko. W związku z tym, że (jak się okazało) tak unikatowa struktura pojawiała się zarówno w anonimach, jak i listach znanego autorstwa, można było z dużym prawdopodobieństwem ocenić, iż autorem obu jest ten sam człowiek.

W celu jeszcze lepszego zobrazowania unikatowego charakteru badanych fraz analizę często rozpoczyna się od wpisania w przeglądarkę skróconej wersji badanego wyrażenia, która ma najmniej charakterystyczną formę, a następnie pojedynczo dodaje się każde kolejne występujące w badanej frazie słowo, z założeniem stopniowego zmniejszania się liczby „trafień”, aż ostatecznie wpisując w pole wyszukiwania pełną badaną frazę, uzyskuje się wyniki bliskie zera, co świadczy o niepowtarzalności analizowanej struktury:

<i>String</i>	<i>Instances</i>
I picked	1,060,000
I picked something	780
I picked something up	362
I picked something up like	1
I picked something up like an	0
an ornament	73,700
like an ornament	896
something like an ornament	2
I asked	2,170,000
I asked her	284,000
I asked her if	86,000
I asked her if I	10,400
I asked her if I could	7,770
I asked her if I could carry	7
I asked her if I could carry her	4
I asked her if I could carry her bags	0

Ryc. 5. Przykład uzyskania dowodu unikatowego charakteru frazy dzięki wynikom zaczerpniętym z przeglądarki internetowej

Źródło: J. Olsson, *Word Crime...*, op. cit., s. 141.

Podkreśla się jednak, że wyniki uzyskane dzięki wskazaniom wyszukiwarek nie mogą występować jako jedyny dowód autorstwa. Powodem tego są przede wszystkim następujące w bardzo krótkim czasie istotne zmiany otrzymanywanych wyników w zakresie badania tej samej cechy językowej oraz fakt, iż czas wyszukiwania internetowego jest uzależniony od aktualnego obciążenia serwera, a więc zwykle otrzymane w przeglądarce wyniki wyszukiwania

nie uwzględniają wszystkich ogólnie dostępnych zasobów. Z tego też powodu w różnych przeglądarkach często uzyskuje się odmienne wyniki³⁶. Pozwala to wysnuć wniosek, iż takich internetowych analiz nie można uznawać za całkowicie wiarygodne i dające pełny obraz użycia języka w społeczeństwie. Doskonale jednak spisują się one jako ewentualne uzupełnienie innych metod lub też wzmocnienie wiarygodności wyników uzyskanych innym sposobem. Z uwagi na to wątek wyników uzyskanych z przeglądarek internetowych pojawia się w większości współczesnych analiz lingwistycznych. Jako przykład sukcesu takiej łączonej analizy warto przytoczyć sprawę zniknięcia kobiety, która po zaginięciu miała wysłać do swojego partnera liczne SMS-y z wyjaśnieniami, gdzie się aktualnie znajduje:

Stopping at jills, back later need to sort my
head out
Tell kids not to worry. sorting my life out. be
in touch to get some things

Ryc. 6. Fragmenty SMS-ów wysyłanych przez zaginioną kobietę

Źródło: J. Olsson, *Word Crime...*, op. cit., s. 141.

W treści tych wiadomości eksperci zidentyfikowali dwa bardzo charakterystyczne wyrażenia: „sort my head out” oraz „sort my life out”, które okazały się na tyle rzadkie, że w korpusie liczącym 100 mln słów został ujawniony tylko jeden przykład użycia struktury „sort my life out”, a struktury „sort my head out” w ogóle nie znaleziono. Natomiast wyniki uzyskane w wyszukiwarce Google wykazały 23 tys. przypadków użycia „sort my life out”, tylko 600 przypadków „sort my head out” oraz jedynie 17 przypadków współwystępowania tych struktur w jednym zdaniu, co pozwalało określić je jako skrajnie rzadkie. Tym bardziej zastanawiający był fakt, iż takich samych wyrażen i w tej samej kolejności, w jakiej pojawiły się w osobnych SMS-ach, użył kochanek zaginionej kobiety podczas przesłuchania:

She was on heavy medication and she said when she'd got her head
sorted out and sorted her life out then it would happen.

Ryc. 7. Fragment przesłuchania kochanka zaginionej

Źródło: J. Olsson, *Word Crime...*, op. cit., s. 141.

³⁶ J. Olsson, *Word Crime...*, op. cit., s. 140.

To oraz inne zaobserwowane podobieństwa pozwalały wysnuć wniosek, iż to podejrzany mężczyzna był autorem wiadomości, które miała napisać zaginiona kobieta. Tak doskonale przeprowadzona analiza lingwistyczna przyczyniła się do ustalenia, iż poszukiwana kobieta została zamordowana przez podejrzanego, co doprowadziło ostatecznie do skazania go na dożywotnie pozbawienie wolności³⁷.

Kolejną grupą stosowanych w lingwistyce kryminalistycznej metod badawczych są analizy jakościowe opierające się przede wszystkim na identyfikacji użytych w tekście form językowych, sposobu ich użycia i próbie odpowiedzi na pytanie, dlaczego autor wybrał właśnie te spośród wszystkich dostępnych środków językowych. Językoznawca w procesie takiej analizy musi odnaleźć wszystkie charakterystyczne elementy danego języka osobniczego, opisać je, odpowiednio sklasyfikować i przypisać im odpowiednią wartość identyfikacyjną³⁸. Zatem w przypadku tych metod nie można się ograniczać do prostego zliczania częstotliwości obecnych w tekście cech, ale analizować je w ścisłym związku z tematem i kontekstem wypowiedzi, konkretną relacją nadawca–odbiorca, czasem i miejscem stworzenia komunikatu³⁹. Najważniejszym zadaniem tych analiz jest odpowiednie scharakteryzowanie zidentyfikowanego w badanym tekście języka osobniczego, zwanego *idiolektem* (języka właściwego tylko dla pojedynczego, konkretnego użytkownika w danym momencie jego rozwoju⁴⁰), a w tym odpowiednia interpretacja znaczenia nieuświadamianych nawyków językowych oraz charakterystycznych słów, struktur zdaniowych i błędów językowych. Z uwagi na tę wieloaspektowość badań jakościowych każdy ekspert językoznawstwa ma wypracowane w tym zakresie własne podejście i metody analiz, które uważa za najbardziej miarodajne do konkretnej sytuacji. Przy tym nie jest możliwe zebranie próbek językowych uwzględniających wszystkie warianty użycia języka osobniczego w każdej możliwej sytuacji przez każdego posługującego się danym językiem człowieka. Co więcej, każdy przejaw użycia języka niesie w sobie co najmniej kilkadziesiąt cech, które można badać. Toteż każda jakościowa analiza językowa będzie się charakteryzować nieuchronnym, mniejszym lub większym subiektywizmem w wyborze badanych cech i w pewnym sensie ich autorską interpretacją. Pełne ujednoczenie metod takich badań, na kształt konstruowania uniwersalnych algorytmów, jak w badaniach ilościowych, jest zaś niemożliwe. Podejmuje się jednak próby sporządzania wykazów cech, które można znaleźć w two-

³⁷ Ibidem, s. 6.

³⁸ G. McMenemy, op. cit., s. 143.

³⁹ M. Zaśko-Zielińska, op. cit., s. 24.

⁴⁰ K. Polański, op. cit., s. 243.

rach językowych, wraz z przypisaniem im odpowiedniej wartości identyfikacyjnej⁴¹. Poniżej przedstawiono tabelę, która w założeniu ma ujednoczyć metody jakościowe przez wskazanie, jakie przejawy języka i w jakiej kolejności powinny się poddawać badaniom.

<i>Feature</i>	<i>No.</i>	<i>Comment</i>	<i>Assessment</i>
Types of grammatical structure	1	See Chaski (1998, 2001b, 2005). A rare, unusual or erroneous structure across known and questioned. Knowledge of generative grammar helpful.	Significance depends on number and type of matches.
Punctuation	2	Unusual punctuation habits or punctuation errors can also be significant. Consider combining punctuation analysis with syntax.	Can be significant, but not necessarily on its own.
Idiom	3	Curious phrasing can be significant. Research its markedness wherever possible.	Depends on number and type of matches.
Spelling	4	Spelling errors, especially rare spelling errors. Look for dialectal spellings, such as US vs UK spellings in English.	As above, but research degree of rarity.
Document layout	5	This can be of interest if the layout style is unusual. Beware, however, of cultural styles of layout.	Caution required, e.g. US vs UK vs European styles.

Ryc. 8. Tabela porządkująca cechy językowe w kolejności od najważniejszej do najmniej istotnej w badaniach identyfikacyjnych

Źródło: J. Olsson, *Forensic Linguistics...*, op. cit.

Należy zauważyć, iż zarówno metody jakościowe, jak i ilościowe stosowane samodzielnie mogą narażać ekspertyzy na błędy. Metody ilościowe, zazwyczaj skomputeryzowane, mogą dostarczyć obiektywne wyniki w bardzo krótkim czasie, jednak nakazują przy tym stosowanie skomplikowanych, niezrozumiałych dla przeciętnego człowieka algorytmów i zwykle badają pojedyncze cechy wypowiedzi w zupełnym oderwaniu od innych występujących

⁴¹ Również na gruncie polskim stworzono podobny katalog cech językowo-treściowych przyjęty przez Jesienną Szkołę Empirycznych Badań Pisma Ręcznego, w: A. Feluś, *Z zagadnień językoznawstwa kryminalistycznego*, Wydawnictwo Naukowe WSPiA im. Mieszka I w Poznaniu, Poznań 2011, s. 141–143.

w tekście. Ponadto nie biorą pod uwagę założeń teorii zmienności języka w zależności od sytuacji komunikacyjnej i kontekstu, co sprawia, że błędne dopasowanie autorstwa zdarza się w nich bardzo często⁴². Natomiast metody jakościowe, jako w dużej mierze polegające na założeniach samej teorii językoznawstwa, wydają się bardziej skuteczne w ustalaniu autorstwa, jednak także napotykają wiele trudności, takich jak na przykład stosunkowo długi czas potrzebny na ich przygotowanie, pełne uzależnienie wyniku od doświadczenia i wiedzy eksperta, a więc subiektywizm i wielokrotnie zarzucany intuicjonizm w dobieraniu badanych właściwości językowych, oraz częste analizowanie pojedynczych środków językowych w oderwaniu od komunikatu językowego jako całości. Wobec tego ani metody jakościowe, ani ilościowe stosowane osobno nie mogą dać zupełnej pewności wyniku analizy. Z jednej strony zauważa się także, że bez gruntownego przygotowania w zakresie teorii językoznawstwa trudno poprawnie zinterpretować wyniki analiz ilościowych, a z drugiej strony analizy jakościowe nieoparte wskazaniem choćby analiz korpusowych mogą doprowadzić do błędnych ocen.

W związku z tą niepewnością i niesprawdzalnością wyników w Stanach Zjednoczonych wielokrotnie zarzuca się analizom lingwistycznym niespełnianie standardów Dauberta⁴³, a tym samym odmawia się takim ekspertyzom przymiotu dowodu w sprawie, co częstokroć może prowadzić do skomplikowania i przeciągania w czasie procesów sądowych, które analizy językowe mogłyby z łatwością rozstrzygnąć. Dlatego też, jak postuluje większość przedstawicieli językoznawstwa kryminalistycznego, analizy lingwistyczne należałoby przeprowadzać, łącząc kilka metod w obrębie jednej ekspertyzy. Obszerne, wielowątkowe badania, poparte wyliczeniami, wykresami, tabelami, mogą przyczynić się do skorygowania ewentualnego marginesu błędu, a także okazać się bardziej czytelne i zarazem bardziej akceptowalne jako dowód autorstwa przez wymiar sprawiedliwości. W związku z tym twórcy programu ALIAS słusznie proponują trzy rodzaje ekspertyz dokonywanych z użyciem tego programu – oparte tylko na stylometrii wspomaganą komputerowo, łączące zautomatyzowane analizy stylometryczne z manualnym opracowaniem ich wyników przez eksperta oraz łączące w sobie analizę ilościową z analizą jako-

⁴² Obszerna krytyka wielu metod ilościowych została przedstawiona przez T. Granta w jego pracy doktorskiej: T. Grant, *Authorship Attribution in a Forensic Context*, University of Birmingham, listopad 2004.

⁴³ Ekspertyza taka musi być: oparta na przetestowanej już procedurze, powszechnie zaakceptowana w środowisku naukowym, opisana w publikacjach naukowych, opatrzona znanym wskaźnikiem błędu, w: M. Coulthard, A. Johnson, *The Routledge Handbook...*, op. cit., s. 479.

ściową opracowaną manualnie przez eksperta językoznawcę – których wybór uzależniają od rodzaju problemu postawionego w zleceniu ekspertyzy.

Powracając zaś do kwestii rozwoju językoznawstwa kryminalistycznego w Polsce, należy mieć nadzieję, iż z biegiem czasu, dzięki czerpaniu z teorii i praktyki wypracowanej w tym zakresie w krajach anglojęzycznych, dziedzina ta nabierze równie szybkiego tempa rozwoju oraz uzyska status i popularność równą tym, jakie ma obecnie w Wielkiej Brytanii i Stanach Zjednoczonych. Niezależnie od tego, jakie podejście do językoznawstwa kryminalistycznego mają obecnie przedstawiciele polskiego wymiaru sprawiedliwości, a przede wszystkim organów ścigania, międzynarodowa praktyka pokazuje, że lingwistyka kryminalistyczna w wielu sprawach może być bardzo pomocnym narzędziem.

Streszczenie

Lingwistyka kryminalistyczna jako dziedzina kryminalistyki funkcjonuje w praktyce już od kilkudziesięciu lat, jednak po dziś dzień wielu autorów określa ją jako nową. Szczególnie na gruncie polskim daje się zauważyć brak zainteresowania możliwościami, jakie daje analiza językowa, nie tylko w procesie typowania nieznanego autora tekstu (profilowanie językowe), ale także udowadniania winy, np. przez przypisanie autorstwa lub ustalenie intencji nadawcy komunikatu. W artykule skupiono się na funkcjonowaniu językoznawstwa kryminalistycznego w krajach anglojęzycznych z uwagi na fakt, iż to właśnie w tych krajach dziedzina ta rozwija się najszybciej. Przedstawiono pokrótce najważniejsze zakresy zastosowania analizy językowej w procesie sprawowania wymiaru sprawiedliwości oraz omówiono najczęściej stosowane w tej dziedzinie lingwistyczne metody badawcze, opatrując je praktycznymi przykładami wykorzystania poszczególnych rodzajów analiz.

Słowa kluczowe: językoznawstwo kryminalistyczne, lingwistyka kryminalistyczna, lingwistyka korpusowa, analiza lingwistyczna, profilowanie językowe, metody ilościowe, metody jakościowe, stylometria, idiolekt, anonimy

Summary

Forensic linguistics has been practiced as a discipline of forensic science for decades, however, many authors still describe it as a new domain. The lack of interest in possibilities provided by linguistic analysis, not only in the process of identification of an unknown author of a text (linguistic profiling) but also in the process of proving a person guilty, for example, by means of ascribing authorship to a person or establishing of the intentions of the author of a text, has been especially clear in Poland. The article focuses on the functioning of forensic linguistics in English speaking countries because in these countries the discipline is developing faster than anywhere

else. It presents a short description of the main scope of application of linguistic analysis in the process of administration of justice and discusses the linguistic research methods which are applied the most frequently in this area, providing examples of practical application of specific types of analysis.

Keywords: forensic linguistics, corpus linguistics, linguistic analysis, forensic linguistics profiling, quantitative methods, qualitative methods, stylometry, idiolect, anonymous letters