

УДК 811.112'2:81'33

Kovbasyuk L. A.

DEUTSCHE KORPORA DER INTERNETBASIERTEN UND MOBILEN KOMMUNIKATION

Статтю присвячено німецьким корпусам інтернет-комунікації та мобільної комунікації. Визначено поняття «інтернет-комунікації», «мобільна комунікація», «корпус». Розглянуто теоретичні засади корпусної лінгвістики. Висвітлено найвідоміші корпуси німецької мови інтернет-комунікації та мобільної комунікації, окреслено їхні характерні ознаки.

Ключові слова: німецька мова, корпус, інтернет-комунікація, мобільна комунікація.

Ковбасюк Л. А. Немецкие корпусы интернет-коммуникации и мобильной коммуникации. – Статья.

Статья посвящена немецким корпусам интернет-коммуникации и мобильной коммуникации. Определены понятия «интернет-коммуникация», «мобильная коммуникация», «корпус». Показаны теоретические основы корпусной лингвистики. Представлены известные немецкие корпусы интернет-коммуникации и мобильной коммуникации. Очерчены их характерные особенности.

Ключевые слова: немецкий язык, корпус, интернет-коммуникация, мобильная коммуникация.

Kovbasyuk L. A. German corpora of Computer-Mediated Communication and of Mobile Communication. – Article.

The article deals with German corpora of Computer-Mediated Communication and of Mobile Communication. The terms “Computer-Mediated Communication”, “Mobile Communication”, “corpus” are defined. The theoretical background of the corpus linguistics is determined. The most important corpora of Computer-Mediated Communication and Mobile Communication are described and their features are analyzed.

Key words: German, text corpora, Computer-Mediated Communication, Mobile Communication.

Zu den wichtigsten Schwerpunkten der modernen Linguistik sowohl in der Ukraine als auch in Deutschland gehören derzeit korpusbasierte Untersuchungen der Gegenwartssprache. Das Interesse der theoretischen Linguistik an die korpuslinguistische Methodik nimmt rapide zu. “Die Häufigkeitsverteilung von sprachlichen Phänomenen findet neben den traditionell empirisch arbeitenden Disziplinen wie Lexikographie, Sprachdidaktik oder Entwicklungspsychologie auch in den aktuellen Arbeiten der kognitiven und der theoretischen Linguistik zunehmend Beachtung. Rasante Fortschritte der Mathematik im Bereich strukturentdeckender Transformationen gestatten es der Korpuslinguistik, immer komplexere sprachliche Phänomene quantitativen Untersuchungen zu unterziehen” [8].

Die Korpuslinguistik ist eine ziemlich junge Sprachdisziplin, die sich seit den 1980er Jahren im englischsprachigen Raum etabliert hat und sich später in vielen Ländern entwickelt hat. Die deutsche Korpuslinguistik als Sprachwissenschaft haben solche Linguisten wie N. Bubenhofer, R. Köhler, L. Lemnitzer, W. Teubert u.a. wesentlich entwickelt [5; 13; 17; 25]. Als Ergebnisse der jahrelangen Arbeit im Bereich der Korpuslinguistik dienen folgende renommierte deutsche Korpora, wie “DeReKo. Das Deutsche Referenzkorpus” [8], “DWDS. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart” u.a. [7]. Korpusbasierte Studien werden in der deutschen Phraseologie [12; 24], im DaF-Unterricht [26] usw. durchgeführt. Aber was Korpora der internetbasierten und der mobilen Kommunikation betrifft, sind sie sowie in der ukrainischen als auch in der deutschen Linguistik unseres Achtens unterrepräsentiert, was die *Aktualität* und *Neuigkeit* des ausgewählten Themas bestätigt.

Das *Ziel* dieses Beitrags besteht darin, die aktuellen deutschen Korpora der internetbasierten und der mobilen Kommunikation festzustellen und sie zu charakterisieren. Welche charakteristischen Merkmale die ausgewählten deutschen Korpora aufweisen ist *Gegenstand* der anschließenden Überlegungen.

Zu den *Aufgaben* des Beitrags gehören: 1) Definition der Termini internetbasierte Kommunikation, mobile Kommunikation, Korpus, 2) Systematisierung der Grundrisse von linguistischen Korpora, 3) Analyse der deutschen Korpora computervermittelter und mobiler Kommunikation.

Die Besonderheiten des Sprachgebrauchs in der Kommunikation per Neue Medien, d.h. solche Medien, die Daten in digitaler Form übermitteln oder auf Daten in digitaler Form zugreifen, und die Analyse ihrer Merkmale, sind Gegenstand der deutschen linguistischen Untersuchungen seit Mitte der 90er Jahren des vorigen Jahrhunderts. Es werden zwei Kommunikationstypen unterschieden: internetbasierte (computervermittelte, computerunterstützte, digitale) und mobile Kommunikation.

Internetbasierte Kommunikation umfasst solche Kommunikationsformen wie E-Mail, Chats, ICQ, Blogs, Instant Messaging, Online-Foren ect., für deren Vermittlung Computernetze (vor allem das Internet) und ein technisches Gerät (Computer/Laptop/iPad) eine notwendige Bedingung darstellen. Internetbasierte Kommunikation gehört zu den wichtigsten Notwendigkeiten des gegenwärtigen menschlichen Alltags. Über die Internet-Kommunikation haben die Leute unabhängig vom Alter eine grenzlose Möglichkeit bekommen, einander verschiedene Informationen, Wünsche, Gedanken weltweit mitzuteilen, obwohl sie sich nicht in einer face-to-face Situation gegenüberstehen. Die Internet-Kommunikation wird aus verschiedenen Sichten sowohl von bekannten

ausländischen als auch von ukrainischen Linguisten untersucht. Innerhalb der linguistischen Disziplin haben sich vor allem C. Durscheidt [11], L. Kovbasyuk [16], L. Kompanzeva [2], P. Schlobinski [19] u.a. mit den sprachlichen Besonderheiten der Internet-Kommunikation beschäftigt.

Mobile Kommunikation (auch SMS-Kommunikation genannt) umfasst solche Kommunikationsformen wie SMS-Berichte, Instant Messaging (Telegram, Viber, WhatsApp) ect., für deren Vermittlung ein mobiles technisches Gerät (Handy/Smartphone/iPhone) eine notwendige Bedingung darstellt. Ein Smartphone ist seit langem nicht nur zum Telefonieren da. Es ist ein Medium mobiler schriftlicher Kommunikation, Plattform für Fotos, Videos, Spiele, Musik, Rundfunk, Online TV, und Internet. Der Short Message Service (SMS) erlaubt kurze Berichte mit einem Umfang bis zu 160 Zeichen über das Tastenfeld des Handys zu senden. Das Instant Messaging ist eine quasi-synchrone Form der mobilen Kommunikation. Das entsprechende Programm stellt fest, wer aus der Familie, der Freunde u.a. gerade online ist, mit wem man also quasi in Echtzeit schnell kommunizieren kann. Die Beiträge folgen unmittelbar aufeinander, bei bestimmten Software-Programmen ist es sogar möglich, dass die Beteiligten zusehen können, wie der andere schreibt, dass die Kommunikation also synchron verläuft [10].

Diese Art der Kommunikation zieht auf sich die Aufmerksamkeit von vielen Sprachforschern. Innerhalb der linguistischen Disziplin haben sich L. Kovbasyuk [1], P. Schlobinski [20] und C. Schmidt [21] mit den sprachlichen Besonderheiten der SMS-Kommunikation beschäftigt. Neuere Publikation von K. König [14] beschäftigt sich mit der Dialogkonstruktion und Sequenzmuster von SMS und WhatsApp-Berichten.

Sowohl internetgestützte als auch mobile Kommunikation dienen nicht nur als Kommunikationsmittel, sondern auch als Gegenstand von vielen Studien aus verschiedener Sichte, besonders verbreitet und aktuell sind zurzeit korpuslinguistische Methoden, die auf verschiedene kleine und große Korpora basiert sind. In der Korpuslinguistik wird der Ausdruck Korpus so definiert [17, s. 7]: Ein Korpus ist eine Sammlung schriftlicher oder gesprochener Äußerungen. Die Daten des Korpus sind typischerweise digitalisiert, d.h. auf Rechnern gespeichert und maschinenlesbar. Die Bestandteile des Korpus, die Texte, bestehen aus den Daten selbst sowie möglicherweise aus Metadaten, die diese Daten beschreiben, und aus linguistischen Annotationen, die diesen Daten zugeordnet sind. Es sei betont, dass da die meisten Korpora hauptsächlich aus geschriebenen Texten erstellt sind, ist ein Nebeneffekt der Korpuslinguistik eine automatische Fokussierung auf geschriebene Sprache. Das heißt, obwohl Korpusstudien auf Sprachgebrauch basieren, haben wir es in den meisten Fällen nicht mit spontan produzierter Sprache zu tun [24, s. 74].

Also, die Korpora sind strukturiert und mit einem Korpusmanager versehen. Korpusmanager sind aufwendige Text-Analyse-Systeme, die komplizierte linguistische Recherchen ermöglichen, wobei die Suche nach sprachlichen Elementen in Form einer Abfrage erfolgt. Zu den üblichsten Funktionen der Korpusmanager, die in den meisten Korpora vorkommen und die die Interpretation der gewonnenen Daten erleichtern, gehören: positive und negative Filter, absolute und relative Frequenzanalyse mit Häufigkeitsklassen von 0 bis 13 und Kollokationsanalysen, d.h. die Analyse des gemeinsamen Vorkommens zweier Wörter in einer strukturell interessanten Einheit. Zu den wichtigsten Anforderungen an die Suchmöglichkeiten in einem Sprachkorpus gehören unter anderem: 1) Abfragen nach Wörtern, Wortteilen und Wortverbindungen und Satzelementen, 2) statistische Angaben über die Suchergebnisse, 3) Angaben über die Größe und Zusammenstellung des Korpus. Als Resultat einer Korpusabfrage erscheinen auf dem Bildschirm Konkordanzen, d.h. "eine Sammlung von Kotexten eines bestimmten Schlüsselworts. Kotexte einer bestimmten Länge (von Buchstaben, Wörtern oder Sätzen) um ein Schlüsselwort herum werden aus einem Korpus extrahiert und meist mit dem Schlüsselwort im Zentrum angeordnet" [17, s. 196]. Zu diesen Konkordanzen gibt es auch Angaben über die Texte, aus denen sie stammen. Zu den weiteren wichtigen Merkmalen der Korpora gehören Annotationen, d.h. zusätzliche Informationen zu den Korpus-texten (äußere Annotation) und zu den einzelnen Wörtern im Korpus (innere Annotation). Äußere Annotationen enthalten bibliographische Angaben zum Text und zu seinem Ursprung, wichtige Informationen über Textsorte, stilistische Zuordnung ect. Innere Annotationen beziehen sich hauptsächlich auf das Wort im Text, auf seine morphosyntaktischen Eigenschaften (Tagging), syntaktisch-semantischen Rollen (Parsing) und pragmatische Merkmale (im Sinne von Topik und Fokus) [17, s. 63–84].

Zu den wichtigsten linguistischen Anwendungsgebieten von Korpora gehören: 1) die Erforschung von sprachlichen Strukturen und Varietäten; 2) Wortbildung; 3) Orthographie; 4) Lexikographie; 5) die Erstellung von Grammatiken; 6) Fremdspracherwerb und -vermittlung; 7) die Computerlinguistik [17, s. 157–193].

In der deutschen Germanistik gibt es die Meinung [4, s. 346–347], dass sich die gegenwärtig existierenden Korpora anhand zweier Kriterien in Typen einteilen lassen. Zunächst unterscheidet man solche Korpora, die als empirische Basis für die Fragestellungen eines bestimmten Projekts zusammengestellt wurden, und Korpora, die ohne direkten Bezug zu einem bestimmten Projekt, sondern vielmehr projektneutral als Datenbasis für die Untersuchung verschiedener möglicher Fragestellungen aufgebaut wurden. Das 2. Kriterium ist die Vorstrukturierung der akquirierten

Daten in Bezug auf anvisierte oder mögliche Fragestellungen und Untersuchungsdesigns. Hier gibt es aufbereitete Korpora, bei deren Erarbeitung die Daten für Recherchezwecke Aufbereitungsprozessen unterzogen wurden, und Rohdatenkorpora, deren Daten in dem Zustand belassen wurden, in welchem sie ursprünglich aus dem Internet akquiriert wurden. Aufgrund dieser Kriterien werden 4 Typen von Korpora: 1) projektbezogene Rohdatenkorpora, 2) projektbezogene aufbereitete Korpora, 3) projektneutrale Rohdatenkorpora und 4) projektneutrale aufbereitete Korpora. Die Typologie der Korpora kann auch auf folgenden Kriterien basieren: 1) Funktionalität, 2) Sprachenauswahl, 3) Medium, 4) Annotation, 5) Größe, 6) Persistenz, 7) Sprachbezug und 8) Verfügbarkeit [17, s. 137–142].

Da das Forschungsfeld internetbasierter und mobiler Kommunikation vergleichsweise neu ist, sind Sprachdaten aus Formen digitaler Kommunikation und mobiler Kommunikation bislang nicht in großen ausgewogenen Korpusansammlungen repräsentiert. An den theoretischen und methodologischen Grundlagen der datengestützten Analyse internetbasierter Kommunikation arbeiteten in der BRD 38 Mitglieder des neuen wissenschaftlichen Netzwerks Empirische Erforschung internetbasierter Kommunikation (Empirikom) [9]. Was die Korpuslandschaft internetbasierter und mobiler Kommunikation betrifft, bestehen derzeit viele offene Fragen, die wichtigsten darunter sind folgende [3, s. 2–3]: 1) rechtliche/ethische Fragen: unter welchen Bedingungen können Sprachdaten aus Genres internetbasierter/mobiler Kommunikation für die wissenschaftliche Nutzung archiviert, für linguistische Analysezwecke aufbereitet, annotiert und als Forschungsressource bereitgestellt werden? 2) Fragen der Strukturrepräsentation/Interoperabilität: welche Formate für die Repräsentation von Textgenres lassen sich sinnvoll für die Strukturbeschreibung von solchen Korpora adaptieren? 3) Fragen der linguistischen Annotation: wie können Werkzeuge für die automatische Sprachverarbeitung für den Umgang mit orthographischen Normabweichungen und spezifischen Stilelementen (z.B. Emoticons, Hashtags ect.) sowie für die Behandlung von Phänomenen der konzeptionellen Mündlichkeit angepasst werden? 4) Fragen der Integration von solchen Korpora in bestehende Korpusinfrastrukturen: wie können sie in existierende Infrastrukturen für die Verwaltung, Bereitstellung und Abfrage von Sprachkorpora integriert werden? Wie lassen sie sich vergleichend mit anderen Typen von Korpora (Textkorpora, Korpora gesprochener Sprache) nutzen und analysieren?

In diesem Beitrag werden im Überblick unseres Achtens die bekanntesten Korpora internetgestützter und mobiler Kommunikation veranschaulicht und analysiert.

Der Sprachgebrauch in der *digitalen Kommunikation* ist in wenigen Spezialkorpora berücksichtigt.

Bekannt sind vor allem *das Dortmunder Chat-Korpus* [6] und *Deutsches Referenzkorpus zur internetbasierten Kommunikation (DeRiK)* [3].

Das Dortmunder Chat-Korpus [6] ist ein annotiertes projektneutrales aufbereitetes Spezialkorpus, das seit 2005 frei (80% aller Korpusdokumente) im Internet zur Verfügung steht. Dieses Korpus ist Resultat der Studie am Lehrstuhl für Linguistik der deutschen Sprache und Sprachdidaktik, das unter der Leitung von Prof. Dr. A. Storrer und Dr. M. Beißwenger am Institut für deutsche Sprache und Literatur der Technischen Universität Dortmund realisiert wurde. Das Korpus enthält mit über 140 000 Chat-Beiträgen bzw. 1,06 Millionen laufenden Wortformen umfangreiches Datenmaterial aus diversen Einsatzformen der Chat-Technologie. Der Bestand reicht von Chats im Hochschulkontext (E-Learning, Online-Zusammenarbeit, kollektive Experten-Interviews) und im Praxisbereich Beratung&Support über Chat-Events im Medienkontext (Chats mit Politikern und Medienakteuren oder begleitend zu TV-Ereignissen) bis hin zu Chats im Freizeitbereich, die in Webchat-Communities stattgefunden haben. Die Korpusdokumente wurden anhand einer XML-Sprache für Recherchezwecke aufbereitet. Das Suchwerkzeug *STACCADO* ermöglicht es, auf einfache Weise nach chat-typischen Elementen wie z.B. Emoticons, Adressierungen, Asterisk-Ausdrücken ect. zu recherchieren, beliebige einfache und komplexe Volltext-Suchanfragen zu formulieren oder statistische Auswertungen zum Kommunikationsaufkommen und zum Beitragsverhalten einzelner Chatter in den Teilkorpora oder in einzelnen Korpusdokumenten zu erzeugen. Das Korpus eignet sich für datengestützte linguistische Analysen und für die Gewinnung von authentischen Beispielen zur Sprachverwendung in Chats im DaF-Unterricht. Derzeit wird das Dortmunder Chatkorpus neu strukturiert und an aktuelle Standards der Repräsentation von Korpora im Kontext der Digital Humanities angepasst und in die CLARIN-D (Common Language Resources and Technology Infrastructure, Deutschland)-Korpusinfrastrukturen integriert.

Deutsches Referenzkorpus zur internetbasierten Kommunikation (DeRiK) [3] ist ein projektbezogenes annotiertes Korpus, das in einer Kooperation der TU Dortmund und der Berlin-Brandenburgischen Akademie der Wissenschaften, Zentrum Sprache, seit 2010 aufgebaut wird. DeRiK ist eine Zusatzkomponente zu DWDS-Korpora. Im Fokus der Erforschung internetbasierter Kommunikation stehen nicht beliebige Webtexte, sondern die sprachlichen Äußerungen in dialogischen Webgenres wie z.B. Online-Foren, Chats, Instant Messaging und in Twitter-Postings, in Kommentaren und Diskussionen in Weblogs, auf Videoplattformen (YouTube) und auf den Profiseiten sozialer Netzwerke (Facebook, MySpace) sowie in multi-modalen Kommunikationsumgebungen wie z. B. Skype. Ein Teil des künftigen Referenzkorpus zur

internetbasierten Kommunikation ist *das Blog-Korpus*. Es besteht aus Beiträgen und Kommentaren, die auf Blogs veröffentlicht worden sind. Diese sind mehrheitlich auf Deutsch und die Betreiber haben die Wiederveröffentlichung der Texte mittels Creative Commons Lizenzen ausdrücklich gefördert. Das Korpus wird regelmäßig aktualisiert [7].

Zu erwähnen ist auch das Korpus *What's up, Deutschland?* [27], das als Ergebnis eines nicht-kommerziellen, universitären Forschungsprojekts entstanden ist. Die Universitäten Leipzig, Dortmund, Dresden, Duisburg-Essen, Hannover, Koblenz-Landau und Mannheim haben im Sommer 2014 den Beschluss gefasst, Leute in ganz Deutschland aufzurufen, im Zeitraum 17.11.2014 – 11.01.2015 ihre WhatsApp-Nachrichten zu spenden. Forschungsfragen sind: 1) Wie werden verschiedene deutsche Dialekte in WhatsApp-Berichten verwendet? 2) Wie werden andere Sprachen in deutschen WhatsApp-Beiträgen verwendet? 3) Wie lang bzw. kurz sind WhatsApp-Nachrichten? 4) Wie sieht die Rechtschreibung in WhatsApp-Berichten aus? 5) Wie werden Emojis verwendet? 6) Welche Wörter bzw. Wendungen treten häufig auf? 7) Wie unterscheiden sich WhatsApp-Beiträgen von SMS-Nachrichten?

Korpora zu Werbeformaten im Internet, in denen unterschiedlichste Werbeformate (Banner-, Domain-, E-Mail-Werbung etc.) für linguistische Analyse aufbereitet sind, kann man in der Rubrik Korpora des Mediensprache-Portals finden. Solche Korpora wie *Mini- und smart-Werbung, Werbung im Internet* ect. können online über eine Datenbankschnittstelle abgefragt werden. Die Suchschnittstelle bietet eine Spezifikation von Suchanfragen u.a. nach Werbeformaten, Produktgruppen, Dateiformaten sowie sprachlichen und gestalterischen Stilmitteln an. Es gibt sowohl PDF- als auch Online-Korpora [15].

Was die Korpora *mobiler Kommunikation* betrifft, sei es betont, dass deutsche Korpora in diesem Bereich kaum vorhanden sind und als besonders repräsentativ nicht gelten können. Sie stehen meistens nicht frei zur Verfügung oder werden derzeit erst aufgebaut und aufbereitet.

Vor allem möchten wir *SMS-Mitteilungen* [22] und *SMS-Mitteilungen. Werbung auf mobilen Endgeräten* [23] von medienprache.net erwähnen. Das Korpus *SMS-Mitteilungen* [22] besteht aus 1500 Short-Messages von Schülern und Studenten aus Osnabrück und Hannover. Im Korpus sind SMS-Berichte von verschiedenen Altersgruppen vorhanden, z. B.: 1) jünger als 12 Jahre, 2) 17–21 Jahre, 3) älter als 30 usw. Außerdem werden die Berichte dem Geschlecht nach

geteilt. Die Belege wurden im Zeitraum 01.01.2000 – 30.09.2001 gesammelt und im PDF-Format gespeichert. Das Korpus hat keine Annotationen und Metadateninformationen.

Das Online-Korpus *SMS-Mitteilungen. Werbung auf mobilen Endgeräten* [23] umfasst eine detaillierte Information ausschließlich der SMS-Werbung: den Werbetext selbst, die Absender-Information (Absender (Telefonnummer), Urheber (Opel, Tchibo, Warsteiner usw.), Datum, Zeit), Stilmittel und Auffälligkeiten (Anglizismen, Begrüßungsworte (Liebe Tchibo Freunde; Hallo), Smileys, mit Fehlern ect.). Die Suche kann man nach bestimmten Wortformen starten.

Seit 01.01.2017 begann die Arbeit an dem *Projekt MoCoDa: Mobile Communication Database* [18], gefördert vom Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen. Gegenstand des Projekts ist der Aufbau einer Datenbank für die kontinuierliche, spendenbasierte Sammlung von Dialogsequenzen aus mobiler Kurznachrichtenkommunikation (WhatsApp, SMS&Co.) als Ressource für die linguistische und sprachdidaktische Forschung, Lehre und für die Förderung sprachlicher Kompetenzen im DaF/DaZ-Kontext. Das Projekt wird an der Universität Duisburg-Essen durchgeführt und von M. Beißwenger, W. Imo, E. Ziegler bis Ende 2017 geleitet. Das Korpus soll in Forschungs- und Lehrprojekten eingesetzt werden, die einerseits Migration und Mehrsprachigkeit zum Gegenstand haben und bei denen es andererseits um die Entwicklung von Materialien und -konzepten für den muttersprachlichen Deutschunterricht, für den Unterricht in den Bereichen Deutsch als Fremdsprache und Deutsch als Zweitsprache sowie im Bereich Sprachliche Integration geht. Alle Daten, die in der Datenbank archiviert und für die spätere Analyse aufbereitet werden, sollen anonymisiert werden. Die Nachrichten müssen so eingetragen/abgeschrieben werden, wie sie im Original versendet bzw. empfangen wurden.

Die durchgeführte Analyse bietet im Überblick eine Veranschaulichung der aktuellen Forschungsbereiche der deutschen Germanisten auf dem Gebiet internetbasierter und mobiler Kommunikation aus der Sicht der Korpuslinguistik. Besonders bekannte und gut ausgebaute deutsche Korpora werden dargestellt und analysiert. Die *Perspektive* der zukünftigen Forschungen liegt in den weiteren Auseinandersetzungen mit Korpora internetbasierter und mobiler Kommunikation aus der Sicht von unterschiedlichen linguistischen Teildisziplinen der modernen Sprachwissenschaft.

Literaturverzeichnis

1. Ковбасюк Л. Instant Messaging vom Smartphone aus linguistischer Sicht / Л. Ковбасюк // Науковий вісник Херсонського державного університету. Серія «Лінгвістика». – 2017. – Вип. 27. – С. 151–156.
2. Компанцева Л. Интернет-коммуникация: когнитивно-прагматичний та лінгвокультурологічний аспекти : автореф. дис. ... докт. філол. наук : спец. 10.02.02 «Російська мова» / Л. Компанцева. – К., 2007. – 36 с.

3. Beißwenger M. Aufbau eines Referenzkorpus zur deutschsprachigen internetbasierten Kommunikation als Zusatzkomponente für die Korpora im Projekt «Digitales Wörterbuch der deutschen Sprache» (DWDS) / M. Beißwenger, L. Lemnitzer // *Journal for Language Technology and Computational Linguistics*. – 2013. – № 26 (2). – S. 1–22.
4. Beißwenger M. Corpora zur computervermittelten (internetbasierten) Kommunikation / M. Beißwenger // *Zeitschrift für germanistische Linguistik*. – 2007. – № 35 (3). – S. 346–375.
5. Bubenhofer N. Sprachgebrauchsmuster: Korpuslinguistik als Methode der Diskurs- und Kulturanalyse / N. Bubenhofer. – Berlin : de Gruyter, 2009. – 388 s.
6. Das Dortmunder Chat-Korpus [Elektronische Ressource]. – Zugang : <http://www.chatkorpus.tu-dortmund.de/>.
7. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart [Elektronische Ressource]. – Zugang : <http://www.dwds.de>.
8. DeReKo. Das Deutsche Referenzkorpus [Elektronische Ressource]. – Zugang : <http://www1.ids-mannheim.de/kl/projekte/korpora/>.
9. DFG-Netzwerk Empirikom: Empirische Erforschung internetbasierter Kommunikation [Elektronische Ressource]. – Zugang : <https://sites.google.com/view/empirikom/startseite>.
10. Dürscheid C. Medien, Kommunikationsformen, kommunikative Gattungen / C. Dürscheid // *Linguistik online*. – 2005. – Bd. 22. – № 1 [Elektronische Ressource]. – Zugang : <https://bop.unibe.ch/linguistik-online/article/view/752/1283>.
11. Dürscheid C. Merkmale der E-Mail-Kommunikation / C. Dürscheid // *Von «hdl» bis «cul8r». Sprache und Kommunikation in den neuen Medien* / Hrsg. P. Schlobinski. – Mannheim et al. : Dudenverlag, 2006. – S. 104–117.
12. Hümmer Ch. Synonymie bei phraseologischen Einheiten. Eine korpusbasierte Untersuchung / Ch. Hümmer. – Frankfurt am Main : Peter Lang, 2009. – 361 s.
13. Köhler R. Korpuslinguistik – zu wissenschaftstheoretischen Grundlagen und methodologischen Perspektiven / R. Köhler // *Zeitschrift für Computerlinguistik und Sprachtechnologie*. – 2005. – Bd. 20. – Ht. 2. – S. 2–16.
14. König K. Dialogkonstitution und Sequenzmuster in der SMS- und WhatsApp-Kommunikation / K. König // *Travaux neuchâtelois de linguistique*. – 2015. – № 63. – S. 87–107.
15. Korpora bei mediensprache.net [Elektronische Ressource]. – Zugang : <http://www.mediensprache.net/de/medienanalyse/korpora/index.aspx>.
16. Kovbasyuk L. Internetbasierte Kommunikation aus linguistischer Sicht: Weblogs / L. Kovbasyuk // *Наукові записки Ніжинського державного університету ім. М. Гоголя. Серія «Філологія»*. – 2014. – Кн. 2. – С. 99–103.
17. Lemnitzer L. Korpuslinguistik: eine Einführung / L. Lemnitzer, H. Zinsmeister. – Tübingen : Gunter Narr, 2006. – 220 s.
18. Mobile Communication Database [Elektronische Ressource]. – Zugang : https://www.uni-due.de/germanistik/ziegler/projekt_mocoda.php.
19. Schlobinski P. Die Bedeutung digitalisierter Kommunikation für Sprach- und Kommunikationsgesellschaften / P. Schlobinski // *Von «hdl» bis «cul8r». Sprache und Kommunikation in den neuen Medien* / Hrsg. P. Schlobinski. – Mannheim et al. : Dudenverlag, 2006. – S. 26–37.
20. Schlobinski P. Sprachliche Aspekte der SMS-Kommunikation / P. Schlobinski // *Literatur und Medien*. – München : Wermke, 2003. – S. 187–198.
21. Schmidt C. Sprachliche Variation in der SMS-Kommunikation / C. Schmidt // *Von *hdl* bis *cul8r**. Sprache und Kommunikation in den Neuen Medien. – Mannheim : Dudenverlag, 2006. – S. 317–333.
22. SMS-Mitteilungen [Elektronische Ressource]. – Zugang : http://www.mediensprache.net/archiv/korpora/sms_os_h.pdf.
23. SMS-Mitteilungen. Werbung auf mobilen Endgeräten [Elektronische Ressource]. – Zugang : <http://www.mediensprache.net/de/werbesprache/handy/>.
24. Stathi K. Korpusbasierte Analyse der Semantik von Idiomen / K. Stathi // *Linguistik online*. – 2006. – Bd. 27. – № 2. – S. 73–89.
25. Teubert W. Corpus linguistics and lexicography / W. Teubert // *Text Corpora and Multilingual Lexicography*. – Amsterdam/Philadelphia : John Benjamins Publishing Company, 2007 – P. 109–134.
26. Wallner F. Lehren und Lernen mit Korpora im DaF-Unterricht / F. Wallner [Elektronische Ressource]. – Zugang : <https://www.goethe.de/de/spr/mag/20454877.html>.
27. What's up, Deutschland [Elektronische Ressource]. – Zugang : <http://www.whatsup-deutschland.de/wer/mitarbeiter/>.