OPTIMALISASI PENGGUNAAN *CORPUS LINGUISTICS* DALAM PENYUSUNAN KAMUS *AZ-ZIRO'AH* SEBAGAI MEDIA PEMBELAJARAN BAHASA ARAB

Yoke Suryadarma, Alinda Zakiyatul Fakhiroh
Universitas Darussalam Gontor, Ponorogo
yoke.suryadarma@unida.gontor.ac.id

Abstrak: Penelitian dalam pembelajaran Bahasa Arab di era Covid-19 ini terus berlangsung. Diantaranya adalah penelitian mengenai pembuatan kamus dwibahasa Indonesia - Arab dalam bidang agroindustri yang meliputi bidang pertanian, teknologi dan industri. Universitas Darussalam Gontor merupakan universitas yang fokus dalam pengembangan dan peningkatan pembelajaran bahasa Arab. Di dalamnya terdapat 7 Fakultas dan 18 Prodi yang mewajibkan mahasiswanya menulis skripsi dalam bahasa asing terutama bahasa Arab, tidak terkecuali Prodi Teknologi Industri Pertanian (TIP). Bertolak dari sini, maka penyusunan kamus dwibahasa ini menjadi penting dilakukan sebagai media pembelajaran bahasa Arab bagi mahasiswa Prodi TIP UNIDA Gontor ataupun pembelajar bahasa Arab lainya serta bagi perkembangan keilmuan bahasa Arab di Indonesia. Makalah ini merupakan gambaran penyusunan kamus *Az-Ziro'ah* yang menggunakan aplikasi linguistik korpus (*corpus linguistic*) dalam penyusunan kosakata (entri dan lemma) di dalamnya. Pendekatan yang dipakai dalam makalah ini adalah pendekatan kualitatif yang menjelaskan secara deskriptif analisis penggunakan aplikasi linguistik korpus (*corpus linguistic*) tersebut dalam penyusunan kamus Az-Ziro'ah. Sehingga fokus makalah ini sebenarnya adalah teknik penyusunan kamus *Az-Ziro'ah* dengan menggunakan pendekatan linguistik korpus Arab (*Arabic Corpus Linguistic*).

Kata kunci: Linguistik korpus Arab, media pembelajaran bahasa Arab, kamus Az-Ziro'ah, kamus Indonesia - Arab, UNIDA Gontor

PENDAHULUAN

Kamus merupakan salah satu kebutuhan utama bagi para pelajara bahasa Asing. Keberadaan kamus menjadi hal penting dalam menjunjang kebutuhan pembelajaran bahasa asing tersebut. Survei yang pernah dilakukan oleh Hartmann menunjukkan bahwa kamus dwibahasa cenderung menjadi alat yang paling disukai oleh pembelajar bahasa, terutama ketika mencari arti kata, tata bahasa, atau konteks pemakaian kata dalam bahasa sasaran (Budiwiyanto, 2014). Pentingnya keberadaan kamus dwibahasa ini pun diungkapkan oleh Utami bahwa keberadaan kamus dwibahasa menjadi hal yang sangat penting sebagai pendamping dalam dunia akademik, namun pengunaannya masih jarang ditemukan dalam dunia akademik (Utami et al., 2019). Oleh karena itu keberadaan dan penggunaan kamus dwibahasa sangat diperlukan sesuai dengan tingkat kebutuhan penggunaan bahasa itu sendiri.

Dalam kaitannya dengan pembelajaran bahasa Arab, kamus dwibahasa Indonesia-Arab merupakan salah satu unsur penting dalam pembelajaran tersebut, baik untuk guru maupun pelajar bahasa Arab. Hal ini dikuatkan oleh penelitian yang dilakukan oleh Rustandi, bahwa peran kamus dwibahasa diperlukan dalam membantu pembelajar memahami dan berkomunikasi dalam bahasa target, dan kamus dwibahasa dijadikan sumber kedua saat

penutur asli tidak ada (Rustandi, 2016). Sedangkan menurut an-Nasywān sebagaimana yang dikutip oleh Wahab, bahwa kompetensi penggunaan kamus, terutama bagi peserta didik non-Arab, sangat diperlukan dalam proses pembelajaran bahasa Arab, sehingga mereka memiliki kemandirian belajar, terutama dalam penguasaan kosakata dan komponen kebahasaan lainnya dari kamus yang digunakan (Wahab, 2017).

Dari beberapa penelitian diatas terlihat bahwa penyusunan kamus dwibahasa menjadi hal yang terus dilakukan dan dikembangkan, terlebih dalam pembelajaran bahasa Arab yang ada di Indonesia. Salah satunya adalah penyusunan kamus *Az-Ziro'ah* yang merupakan kamus dwibahasa Indonesia - Arab dalam bidang Teknologi Industri Pertanian.

Penyusunan kamus dwibahasa ini berangkat dari permasalahan yang ditemukan yaitu adanya kesulitan bagi mahasiswa Prodi Teknologi Industri Pertanian (TIP Fakultas Sains dan Teknologi dalam pembuatan makalah ilmiah dan tugas akhir (skripsi) berbahasa Arab (Suryadarma et al., 2019), salah satu penyebabnya ialah tidak adanya mahasiswa yang menulis skripsi menggunakan bahasa Arab, sebagian besar dari mereka menulis dengan menggunakan bahasa Inggris yang dinilai lebih mudah dalam penulisan istilah-istilah yang berhubungan dengan bidang ilmu mereka masing-masing. Hal ini terjadi karena belum banyaknya kamus-kamus Indonesia-Arab yang memudahkan mereka dalam menerjemahkan istilah-istilah ilmiah sesuai dengan prodi masing-masing ke dalam bahasa Arab, disebabkan kamus-kamus bahasa Indonesia – Arab atau sebaliknya yang ada saat ini masih bersifat umum, dan belum bersifat khusus pada suatu term tertentu (Akbar, 2015).

Penyusunan kamus dwibahasa *Az-Ziro'ah* ini berbasis linguistik korpus. linguistik Korpus merupakan database bahasa yang diambil sampelnya secara statistik untuk tujuan penyelidikan, deskripsi, aplikasi, dan analisis yang relevan dengan semua cabang linguistik (Dash, 2008). Korpus tersebut berisikan kumpulan data, baik data biasa maupun data digital, dalam bentuk tertulis yang berisi bermacam- macam informasi kebahasaan, mulai dari tataran kata, struktur, makna, dan wacana, yang dapat dimanfaatkan untuk penelitian (Hizbullah et al., 2016). Sedangkan menurut Kridalaksana, korpus merupakan kumpulan ujaran tertulis atau lisan yang dipergunakan untuk menyokong atau menguji hipotesis tentang struktur bahasa (Kridalaksana, 2009). Korpus dihasilkan melalui suatu proses digitalisasi dapat dimanfaatkan dalam berbagai bidang linguistik, seperti morfologi, sintaksis, semantik, leksikologi, maupun bidang kebahasaan lain seperti penerjemahan, kesusastraan, dan sebagainya (Hizbullah et al., 2016). Dari sini terlihat jelas bahwa linguistik korpus dapat dimanfaatkan untuk mengkaji dan melakukan berbagai penelitian bahasa, baik itu teori maupun terapan dalam pembelajaran bahasa Arab (Hizbullah et al., 2020), salah satunya adalah dengan penyusunan kamus dwibahasa.

Dalam makalah ini, peneliti berusaha mendeskripsikan optimalisasi pendekatan linguistik korpus dalam penyusunan kamus dwibahasa Az-*Ziroa'h* dalam bidang agroindustri atau Teknologi Industri Pertanian.

METODE

Penelitian ini merupakan penelitian lapangan jenis deskriptif kualitatif (Sugiyono, 2012). Dimana penelitian ini bertujuan mendeskripsikan penyusunan kamus dwibahasa Indonesia-Arab berbasis linguistik korpus dalam bidang Teknologi Industri Pertanian yang ditujukan bagi mahasiswa prodi Teknologi Industri Pertanian (TIP) UNIDA Gontor sebagai media pembantu pembelajaran bahasa Arab tanpa banyak menyajikan data-data yang bersifat kuantitatif dan menitikberatkan kepada peran Peneliti sebagai instrumen utama penelitian.

Adapun Adapun metode yang digunakan oleh Peneliti dalam penelitian pembuatan kamus (leksikografi) ini, adalah metode penyusunan kamus yang diambil mengambil dari Schierholz menggunakan teori yang dikembangkan oleh Wiegand (Wiegand, 1998), yaitu

melalui lima tahapan sistematis yaitu fase persiapan, pengumpulan data dan bahan, pengolahan data dan bahan atau pemprosesan data, evaluasi materi atau penilaian data dan bahan, dan persiapan pencetakan kamus (Schierholz, 2015).

Terkait dengan sumber data yang dijadikan sumber pengambilan data, peneliti mengambil dari 19 artikel ilmiah yang ada di http://garuda.ristekdikti.go.id/ dengan memasukan kata kunci "agroindustri pertanian" dan juga beberapa karya ilmiah dari https://ejournal.unida.gontor.ac.id/index.php/atj. Selain itu data juga diambil dari beberapa dokumen prodi TIP. Sedangkan untuk pengolahan data, peneliti menggunakan pendeketan linguistik korpus sebagai basis dari penyusunan kamus melalui aplikasi web https://www.sketchengine.eu/ untuk melihat kosakata-kosakata yang digunakan dalam bidang agroindustri atau teknologi industri pertanian melalui kajian *wordlist* (https://www.sketchengine.eu/guide/wordlist-frequency-lists/). Sketchengine sengaja dipilih karena memungkinkan peneliti untuk membangun korpus (*building corpora*) sendiri sesuai dengan bidang yang ingin diteliti.

HASIL TEMUAN DAN DISKUSI HASIL

Setelah mengadakan penelitian dalam penyusunan kamus dwibahasa ini, peneliti menemukan bahwa pendekatan linguistik korpus ini sangat membantu peneliti dalam menyusun kamus dwibahasa Indonesia-Arab. Hal ini diperkuat oleh temuan dash (2008) bahwa corpora bahasa adalah satu-satunya sumber dari mana leksikograf dapat memperoleh informasi dari berbagai jenis untuk disajikan dalam kamus dengan optimal (Dash, 2008). Selain itu, hall ini pun juga dibuktikan dengan penyusunan kamus yang telah dilakukan oleh Azzahra dkk dimana penelitian tersebut berhasil mendapatkan 400 kosakata dalam bidang kedokteran yang diambil dari beberapa situs berita Arab(Azzahra et al., 2020).

Penelitian yang dilakukan oleh Hizbullah juga menunjukan hal yang sama, diantara fungsi pemanfaatan kajian linguistik korpus dalam pembelajaran bahasa Arab adalah proyek penyusunan kamus dalam suatu bidang yang spesifik (Hizbullah et al., 2020). Salah satu penyebabnya adalah karena kajian linguistik korpus dapat mengolah data berdasarkan satu term bidang tertentu (Hizbullah et al., 2020), seperti bidang agroindustri atau teknologi industri pertanian ataupun di bidang lainnya, sehingga sangat sesuai digunakan oleh peneliti bahasa (*linguist*) atapun pembuat kamus (*lexicographers*). Adapun hasil penelitian penyusun kamus ini adalah sebagai berikut :

Pengembangan metode penyusunan berbasis linguistik korpus

Dalam penyusun kamus ini, Peneliti menggunakan metode adalah sebagai berikut :

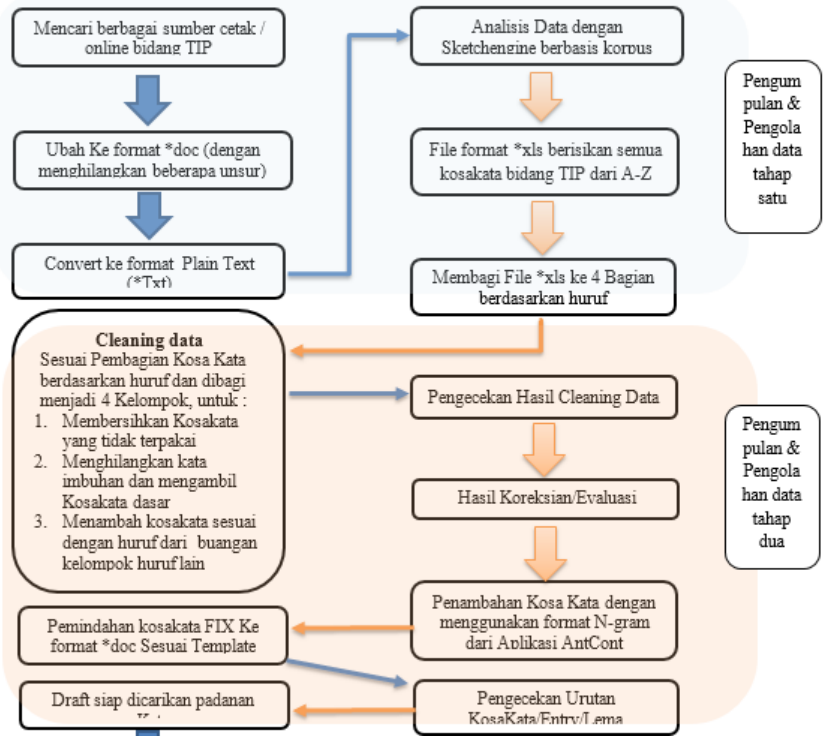
1. Fase Persiapan, meliputi: Perumusan konsep kamus, Pembuatan buku intruksi pembuatan kamus, Pembuatan template draft kamus.
2. Fase Pengumpulan & Pengolahan data, meliputi : Pengumpulan & Pengolahan data tahap satu, Pengumpulan & Pengolahan data tahap dua, dan Pengumpulan & Pengolahan data tahap Tiga.
3. Fase Penyelesaian, meliputi : Penilaian kamus oleh Pakar Leksikografi & pakar Linguistik, Revisi kamus, Pencetakan kamus, dan Pendaftaran HaKI.

Ketiga tahapan ini merupakan metode penyusunan kamus yang peneliti kembangkan berdasarkan teori penyusunan kamus yang dikemukakan oleh Schierholz (Schierholz, 2015) menggunakan teori yang dicetuskan oleh Wiegand (Wiegand, 1998). Adapun bagian yang dikembangakan adalah dalam tahapan pengumpulan dan pengolahan data, dimana peneliti pengembangkan tahapan penyusunan kamus berbasis linguistik korpus. Setidaknya ada beberapa hal yang peneliti lakukan, yaitu :

1. Pengumpulan kosakata-kosakata bidang agroindustri (teknologi, industri dan pertanian) melalui artikel dari media cetak dan media online. Adapun artikel dari media cetak diambil dari dokumentasi prodi TIP meliputi profil prodi, materi perkuliahan, dsb. Sedangkan media online, diambil dari : 19 artikel jurnal ilmiah bidang teknologi industri pertanian online yang ada di http://garuda.ristekdikti.go.id/, salah satunya adalah Jurnal manajemen dan agribisnis milik IPB dan Jurnal Agroindustrial Tecnology Jurnal milik Prodi TIP UNIDA Gontor.
2. Penginputan data tersebut dan penyeleksiannya dalam format *doc dengan membuang semua unsur-unsur tulisan redaksi (tanda baca, angka, dll) yang tidak diperlukan.
3. Mengkonversi file yang sudah fix ke Plain Text (*txt) engan unicode UTF-8.
4. Analisis data berbasis linguistik korpus menggunakan aplikasi Sketchengine melalui : https://www.sketchengine.eu/, menggunakan menu *wordlist* sehingga menghasilkan satu file analisis data yang berformat *xls berisikan semua kosakata bidang TIP berurutan dari Huruf A-Z.
5. Melakukan *data cleaning* dengan perhatian kepada kosakasa dasar, kosakata turunan dan kata berimbuhan, lalu melakukan penambahan lemma/entri berupa kosakata istilah dasar yang belum ada dalam file dengan menggunakan format N-gram dari Aplikasi AntCont (https://www.laurenceanthony.net/software/antconc/).
6. Merevisi data dan memindahkannya ke file dengan format *doc sesuai dengan template yang sudah disiapkan.
7. Draft kamus dalam format *doc siap dicarikan padanan katanya.

Berikut alur kerjanya :

Gambar 1. Pengumpulan & Pengolahan data berbasis linguistik korpus



Dari hasil pengumpulan dan pengolahan data berbasis linguistik korpus tahap pertama didapatkan 4733 kosakata yang berhubungan dengan bidang agroindustri atau teknologi industri pertanian. Data ini didapatkan dari hasil analisis olah data menggunakan kajian *wordlist* dari aplikasi sketchengine. Kemudian setelah dilakukan penyuntingan data, dan *data cleaning* serta penambahan kosakata istilah, di tahap pengumpulan dan pengolahan data tahap kedua, didapati 3548 kosakata terdiri atas kata dasar, dan kata turunan dalam bentuk kata kerja, kata benda, kata sifat, dan sebagainya yang berhubungan dengan bidang agroindsutri/teknologi industri pertanian. 3548 kosakata itulah yang nantinya akan dimasukan ke draft kamus dalam bentuk *doc untuk kemudian dicarikan padanan katanya sesuai dengan buku pedoman penyusun kamus ini.

KESIMPULAN

Hasil penelitian ini menunjukan bahwa linguistik korpus dapat digunakan sebagai pendekatan atau basis dari penyusunan kamus dwibahasa bidang spesifikasi tertentu. Selain itu penyusunan kamus dwibahasa berbasis korpus ini mempunyai hasil yang sangat signifikan dalam membantu penyusun untuk memetakan dan mengelompokan kosakata-kosakata khusus dalam suatu term atau bidang tertentu dengan mudah dan sistemastis. Tahapan yang terpenting dalam penyusun kamus dwibahasa berbasis linguistik korpus ini adalah pemilihan sumber data yang akan diteliti, mekanisme pengolahan data yang terdapat dalam aplikasi pengolah data korpus dan kemampuan peneliti di dalam penyeleksian kosakata atau lemma yang akan dimasukan nantinya ke dalam draft kamus. Ketiga kemampuan tersebut menjadi syarat utama keberhasilakan penyusunan kamus berdasarkan linguistik korpus.

SARAN-SARAN

Walaupun Penyusunan kamus dwibahasa ini telah berhasil dilakukan berbasis linguistik korpus, tetapi peneliti melihat masih ada beberapa kekurangan. Salah satunya adalah ketiadaan pemanfatan korpus paralel dalam penyusunan kamus dwibahasa ini. Pemanfatan korpus pararel ini sangat penting dilakukan untuk memaksimalkan hasil padanan kata dan juga membantu penyusun dalam pemilihan padanan kata yang sesuai. Selain itu penyusunan kamus dwibahasa ini belum bisa dikatakan sempurna, oleh karena itu, penelitian lanjutan masih sangat terbuka bagi peneliti selanjutnya untuk memodifikasi dan mengembangkan serta menyempurnakan hasil penelitian ini.

UCAPAN TERIMAKASIH

Ucapan terimakasih peneliti sampaikan kepada Kemenristek/BRIN melalui Deputi Bidang Penguatan Riset dan Pengembangan (DRPM) atas dana hibah yang diberikan kepada peneliti, juga kepada LPPM UNIDA Gontor dan semua civitas akademika Fakultas Tarbiyah UNIDA Gontor atas atas segala dukungan dan motivasi.

DAFTAR PUSTAKA

Akbar, A. K. (2015). منهج معجم المنور لأحمد ورسون منور ومعجم لمحمود يونس (دراسة وصفية مقارنة). *LISANUDHAD*, *2*(2), 67–82. https://doi.org/10.21111/lisanudhad.v2i2.642

Azzahra, S. F., Hizbullah, N., & Suryaningsih, I. (2020). PENYUSUNAN KAMUS KEDOKTERAN ARAB – INDONESIA DENGAN PENDEKATAN LINGUISTIK KORPUS. *Tsaqofiya : Jurnal Pendidikan Bahasa Dan Sastra Arab*, *2*(2), 60–66.

https://doi.org/10.21154/tsaqofiya.v2i2.37 PENYUSUNAN

Budiwiyanto, A. (2014). *Urgensi penyusunan kamus dwibahasa indonesia inggris untuk penutur bahasa indonesia*. Badan Pengembangan Dan Pembinaan Bahasa Kemendikbud RI. http://www.badanbahasa.kemdikbud.go.id/lamanbahasa/content/urgensi-penyusunan-kamus-dwibahasa-indonesia-inggris-untuk-penutur-bahasa-indonesia

Dash, N. S. (2008). *Corpus Linguistics An Introduction*. Dorling Kindersley.

Hizbullah, N., Arifa, Z., Suryadarma, Y., Hidayat, F., Muhyiddin, L., & Kurnia Firmansyah, E. (2020). Source-Based Arabic Language Learning: a Corpus Linguistic Approach. *Humanities & Social Sciences Reviews*, *8*(3), 940–954. https://doi.org/10.18510/hssr.2020.8398

Hizbullah, N., Rachman, F., & Fauziah, F. (2016). LINGUISTIK KORPUS DALAM KAJIAN DAN PEMBELAJARAN BAHASA ARAB DI INDONESIA. *Konferensi Nasional Bahasa Arab (KONASBARA) II*, 385–393.

Kridalaksana, H. (2009). *Kamus linguistik*. Gramedia Pustaka Utama.

Rustandi, E. (2016). Kamus Dwibahasa Arab–Indonesia Untuk Pembelajar Pemula. *Edusentris, Jurnal Ilmu Pendidikan Dan Pengajaran*, *3*(2), 189–199. https://doi.org/10.17509/edusentris.v3i2.217

Schierholz, S. J. (2015). Methods in Lexicography and Dictionary Research. *Lexikos*, *25*, 323–352. https://doi.org/10.5788/25-1-1302

Sugiyono. (2012). *Metode Penelitian Kuantitatif, kualitatif dan R&D* (16th ed.). CV Alfabeta.

Suryadarma, Y., Abdullah, A. F., & Rini, F. S. (2019). Khaṣāiṣ Qāmūs "Mu'jam Aṣ-Ṣiḥḥah" (Indūnīsī-'Arabī) lī Qism Aṣ-Ṣaydāliyyah wa Qism At-Tagẓiyyah Bi Jāmi'ati Dārussalam Gontor wa Muwāṣafātuhū.pdf. *LISANUDHAD*, *6*(2), 111–127. https://doi.org/http://dx.doi.org/10.21111/lisanudhad.v6i2.3521

Utami, A. D., Hardini, T. I., & Mutiarsih, Y. (2019). Leksikografi Kamus Dwibahasa Prancis-Indonesia Khusus Bidang Bisnis. *Masyarakat Linguistik Indonesia*, *37*(1), 45–60. https://doi.org/10.26499/li.v37i1.86

Wahab, M. A. (2017). Peta Perkembangan Leksikografi Arab Di Indonesia: Studi Kritis Atas Kamus Karya Mahmud Yunus. *Arabi : Journal of Arabic Studies*, *2*(1), 19–36. https://doi.org/10.24865/ajas.v2i1.31

Wiegand, H. E. (1998). *Wörterbuchforschung, Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. Walter de Gruyter.