

Volume 17	No. 2, November 2021	Halaman 177-189
-----------	----------------------	-----------------

**PENGAPLIKASIAN ANTCONC PADA KORPUS BAHASA MELAYU AMBON
(The Application of AntConc on Ambon Malay Language Corpus)**

Nita Handayani Hasan

Magister Ilmu Linguistik Universitas Airlangga

Jalan Airlangga No. 4-6 Surabaya, Indonesia

Pos-el: bontanita00kantorbahasapromal@gmail.com

(Diterima: 22 Juni 2020; Direvisi: 15 Februari 2021; Disetujui: 30 Juni 2021)

Abstract

Development of local language corpus is one way to documenting a local language. Corpus application is one solution to develop a local language corpus. In this paper, researcher using AntConc application to develop Ambon Malay language. In order to develop corpus of Ambon Malay language, this study aims to find out how to use AntConc application in corpus of Ambon Malay language, and to find out the vocabularies that appears on the corpus that has been made. This is a qualitative research which using result of AntConc application. The method in this research is descriptive statistics. The data in this study were sourced from Fuli magazines which containing Ambon Malay language articles. The study found that AntConc application is easy to use. AntConc application starting from entering corpus data in txt form; checking words list; checking word position in the sentence; checking the spread words; checking word use location; checking couples word; and checking the word use combination. The menus in the application can facilitate researcher to analyse corpus source and use it as a dictionary source. The result of the analysis found 1,399 word types and 7,005 word tokens.

Keywords: AntConc, corpus, Ambon Malay language

Abstrak

Pengembangan korpus bahasa daerah merupakan salah satu cara pendokumentasian bahasa daerah. Salah satu cara untuk mengembangkan korpus bahasa daerah yaitu dengan pemanfaatan aplikasi korpus. Aplikasi korpus yang akan digunakan dalam pengembangan korpus bahasa Melayu Ambon pada penelitian ini yaitu AntConc. Agar dapat mengembangkan korpus bahasa Melayu Ambon, maka penelitian ini bertujuan untuk mengetahui bagaimana penerapan penggunaan aplikasi AntConc dalam korpus bahasa Melayu Ambon, dan kosakata-kosakata apa saja yang muncul pada korpus yang telah dibuat. Penelitian ini merupakan penelitian kualitatif yang memanfaatkan hasil analisis aplikasi AntConc terhadap korpus bahasa Melayu Ambon. Metode yang digunakan untuk membaca hasil analisis aplikasi AntConc ialah statistik deskriptif. Data pada penelitian ini bersumber dari majalah-majalah Fuli yang berisi artikel berbahasa Melayu Ambon. Hasil yang diperoleh yaitu aplikasi AntConc merupakan aplikasi yang dapat membantu peneliti untuk menemukan kosakata-kosakata bahasa Melayu Ambon dalam majalah Fuli, serta cara penggunaannya. Cara menggunakan aplikasi AntConc dimulai dari memasukkan data korpus dalam bentuk txt.; mengecek jumlah tipe kata yang terdaftar dalam korpus; mengecek posisi kata dalam kalimat; mengecek penyebaran kata dalam teks; mengecek letak penggunaan kata; mengecek pasangan-pasangan kata; dan mengecek kombinasi penggunaan kata. Menu-menu yang terdapat dalam aplikasi AntConc mempermudah peneliti untuk menganalisis sumber korpus dan memanfaatkannya sebagai sumber penyusunan kamus. Hasil analisis lainnya yaitu ditemukan 1.399 tipe kata, dan 7.005 karakter kata.

Kata-kata kunci: AntConc, korpus, bahasa Melayu Ambon

PENDAHULUAN

Bahasa Melayu Ambon hadir di tengah-tengah masyarakat Maluku sebagai bahasa pemersatu. Keberadaan bahasa Melayu Ambon dianggap sebagai alat komunikasi yang efektif bagi para penutur bahasa daerah. Pada masa penjajahan Belanda, bahasa Melayu Ambon digunakan Belanda untuk menghapus bahasa-bahasa daerah di Maluku. Hal tersebut bertujuan untuk mempermudah Belanda dalam hal berkomunikasi dan berkoordinasi dengan masyarakat pribumi.

Bahasa Melayu Ambon juga berperan dalam hal penyebaran agama. Desa-desa yang telah menganut agama Kristen pada abad ke-18 telah meninggalkan bahasa daerahnya, serta telah menggunakan bahasa Melayu Ambon dalam berkomunikasi (Collins, 2018). Selain berfungsi dalam penyebaran agama, bahasa Melayu Ambon juga digunakan dalam perdagangan dan urusan diplomatik.

Bahasa Melayu Ambon memiliki perbedaan dengan bahasa-bahasa Melayu lainnya. Bahasa Melayu Ambon merupakan bahasa yang muncul dari serapan bahasa Portugis, Belanda, dan bahasa-bahasa daerah di Maluku (Collins, 2018). Kuatnya pengaruh Portugis dan Belanda di pulau Ambon, menjadikan bahasa Portugis dan Belanda berasimilasi dengan bahasa-bahasa daerah di pulau Ambon.

Saat ini, bahasa Melayu Ambon telah menjadi bahasa yang dipilih oleh generasi muda di Maluku. Perubahan kosakata-kosakata bahasa Melayu Ambon juga saat ini terus-menerus terjadi. Hal tersebut dikarenakan masuknya pengaruh bahasa Indonesia dalam bahasa Melayu Ambon. Oleh

karena itu, dibutuhkan langkah-langkah preventif untuk melestarikan keberadaan bahasa Melayu Ambon.

Pelestarian bahasa Melayu Ambon dapat dilakukan dengan penyusunan kamus bahasa Melayu Ambon dan penulisan karya-karya nonilmiah dalam bahasa Melayu Ambon. Kamus bahasa Melayu Ambon yang pernah disusun yaitu “Kamus Bahasa Melayu Ambon—Indonesia” (Takaria, D., 1998) dan Kamus Populer Bahasa Melayu Ambon (ForSMAtu, 2016).

Selain kamus, artikel-artikel berbahasa Melayu Ambon juga merupakan sarana pelestarian bahasa daerah, khususnya bahasa Melayu Ambon. Artikel-artikel yang menggunakan bahasa Melayu Ambon dapat dijadikan contoh penggunaan bahasa Melayu Ambon.

Tulisan-tulisan dalam bahasa Melayu Ambon dapat ditemukan dalam majalah *Fuli*. Majalah *Fuli* merupakan majalah dwibahasa terbitan Kantor Bahasa Maluku yang memuat artikel-artikel nonilmiah yang berhubungan dengan kebudayaan-kebudayaan yang ada di Provinsi Maluku. Artikel-artikel yang dimuat harus ditulis menggunakan dua bahasa, yaitu bahasa Indonesia dan bahasa daerah. Para penulis artikel biasanya berasal dari kalangan budayawan, penulis, dan pakar yang memiliki kemampuan berbahasa Indonesia dan bahasa daerah yang baik. Keberadaan majalah *Fuli* merupakan wadah bagi pelestarian bahasa daerah.

Korpus berbentuk potongan teks yang terdiri atas beberapa sumber. Korpus dapat bersumber dari teks buku, koran, majalah, dll. Korpus bahkan dapat berasal dari tuturan alami yang telah ditranskripsi ke dalam bentuk teks.

Data korpus disimpan dalam bentuk elektronik agar dapat diolah oleh mesin atau program komputer. Pemilihan data korpus dilakukan berdasarkan kriteria tertentu, yaitu komposisi teks dan jenis teks.

Data korpus yang akan dijadikan bahan analisis juga harus memperhatikan ragam bahasanya. Jika kita ingin membuat korpus kosakata-kosakata ilmiah, maka teks-teks yang kita kumpulkan harus berasal dari teks-teks yang berisi hasil-hasil penelitian. Teks-teks tersebut biasanya memuat kata-kata baku dan berasal dari ragam formal.

Data korpus merupakan salah satu objek baru dalam penelitian kebahasaan. Bidang kebahasaan yang memanfaatkan data korpus yaitu leksikografi. Data korpus dapat mempermudah para peneliti bahasa dalam menganalisis penggunaan kosakata.

Keuntungan menggunakan data korpus dalam menganalisis suatu bahasa yaitu dapat memudahkan peneliti dalam mengetahui kata-kata atau frasa yang paling sering muncul dalam suatu bahasa. Korpus juga dapat menunjukkan kata-kata apa saja yang paling sering berpasangan dengan suatu kata. Data korpus juga dapat menunjukkan preposisi apa yang paling sering menempel pada suatu kata. Data korpus memudahkan peneliti mengetahui makna yang terkandung dalam suatu kata berdasarkan konteks penggunaannya. Konteks penggunaan tersebut dapat dimanfaatkan sebagai contoh penggunaan kata.

Penelitian-penelitian yang membahas korpus bahasa yaitu penelitian yang dilakukan oleh Nurjanah (Nurjanah, 2018), memaparkan keragaman kosakata dalam korpus sastra. Penelitian tersebut menemukan empat kata dasar dan 88 kata turunan yang belum terekam dalam KBBI edisi kelima. Penelitian lainnya yaitu

penelitian yang dilakukan oleh Fajri (Fajri, 2018), memanfaatkan teks Indo-Belanda untuk mengembangkan korpus bahasa Indonesia. Hasil pendataan yang dilakukan digunakan untuk memetakan pandangan kaum Indo-Belanda terhadap Indonesia dari bahasa yang digunakan dalam teks. Selain itu, hasil penelitian tersebut dapat digunakan sebagai data sejarah perkembangan bahasa Indonesia. Penelitian yang dilakukan Yuliatwati (Yuliatwati, 2018) membandingkan penggunaan kata perempuan dan wanita dalam korpus *IndonesiaWaC (Sketch Engine)* dan *Ind_mixed_2013 (Leipzig Corpora Collection)*. Penelitian yang dilakukan oleh Henry Sujaini (Sujaini, 2018) mengungkapkan peningkatan akurasi penerjemahan bahasa daerah dengan optimasi korpus paralel. Penelitian tersebut memanfaatkan korpus bahasa Indonesia—Melayu dan Indonesia—Jawa untuk menguji mesin penerjemah berbasis statistik. Penelitian yang dilakukan oleh Hizbullah, dkk. (Hizbullah, Fazlurrahman, & Fauziah, 2016) menggunakan manuskrip-manuskrip Arab yang telah terkumpul untuk dijadikan tujuh tipe korpus bahasa bahasa Arab di Indonesia.

Berdasarkan hasil studi pustaka tersebut diketahui bahwa penelitian-penelitian yang berhubungan dengan pengembangan korpus bahasa daerah masih sulit ditemukan. Oleh karena itu, penulis mencoba memanfaatkan artikel-artikel berbahasa Melayu Ambon dalam Majalah *Fuli* sebagai sumber korpus bahasa Melayu Ambon.

Pembuatan korpus bahasa Melayu Ambon akan dilakukan dengan memanfaatkan aplikasi AntConc 3.4.1 m (Macintosh OS X) yang dikembangkan oleh Laurence Anthony. Aplikasi tersebut akan menunjukkan daftar kata bahasa Melayu Ambon dalam majalah *Fuli*.

Permasalahan yang akan dibahas dalam penelitian ini yaitu bagaimanakah penerapan penggunaan aplikasi AntConc dalam korpus bahasa Melayu Ambon dan kosakata-kosakata apa saja yang muncul pada korpus yang telah dibuat. Untuk menjawab pertanyaan-pertanyaan tersebut, penulis akan mendeskripsikan langkah-langkah yang akan digunakan, sehingga diharapkan pembaca mampu memahami penyusunan korpus, serta mampu menjadi langkah awal penyusunan korpus bahasa Melayu Ambon. Penyusunan korpus bahasa Melayu Ambon juga nantinya dapat dimanfaatkan oleh para pekamus, peneliti, dan pengguna bahasa Melayu Ambon.

LANDASAN TEORI

Kajian linguistik korpus merupakan kajian yang tergolong baru. Saat ini korpus digunakan sebagai salah satu data dalam penyusunan kamus. Pemanfaatan korpus dalam pembuatan kamus sangat bermanfaat bagi para pekamus dalam menemukan kosakata dan penyusunan makna pada kosakata. Oleh karena itu, penentuan korpus yang tepat sebagai sumber data dalam penyusunan kamus sangat penting agar data yang diinginkan dapat diperoleh dengan baik.

Kata korpus berasal dari bahasa Latin, *corpus*, yang berarti badan atau kumpulan (Jones, 2008). Kumpulan yang dimaksud yaitu kumpulan teks dari suatu variasi bahasa yang secara maksimal mampu merepresentasikan suatu kecenderungan atau fenomena dari variasi kalimat (McEnery & Wilson, 2007). Agar mampu melihat variasi bahasa atau fenomena bahasa pada sebuah korpus, maka dibutuhkan kumpulan teks yang berjumlah besar.

Kemunculan korpus dalam linguistik dibagi dalam dua fase, yaitu

fase awal dan modern. Pada fase awal, penyusunan korpus dilakukan secara manual. Para pekamus mengumpulkan bahasa lisan dan tulisan dalam waktu bertahun-tahun agar memperoleh jumlah korpus yang representatif sehingga dapat diolah dalam penyusunan kamus. Pada awal kemunculannya, korpus masih dianggap belum mampu dijadikan dasar untuk mengembangkan teori bahasa. Namun, para linguis yang meyakini pentingnya keberadaan korpus tetap mengembangkan penelitiannya. Penelitian-penelitian yang tetap dilakukan tersebut, saat ini menjadi dasar-dasar pengembangan teori bahasa.

Fase modern keberadaan korpus diikuti dengan pengembangan teknologi komputer. Pemanfaatan teknologi komputer pada korpus menjadikan penelitian yang berhubungan dengan korpus semakin berkembang. Saat ini, peneliti dan pekamus dapat membuat korpus bahasa dengan mudah dan kemudian menganalisisnya. Salah satu aplikasi yang dapat dimanfaatkan untuk membuat dan mengembangkan korpus bahasa yaitu AntConc.

Pada umumnya pemanfaatan korpus dalam linguistik korpus khususnya untuk keperluan leksikografi diawali dengan melihat daftar frekuensi kata (*frequency list*) dalam sebuah program. Melalui daftar frekuensi kata, akan diketahui kata-kata apa saja yang paling sering muncul atau yang paling jarang muncul. (McEnery, Tony; Hardie, 2012:243) mengartikan *frequency list* sebagai daftar semua bentuk dari tipe atau jenis sebuah korpus (misalnya berbentuk semua kata, semua bentuk *part of speech*, semua *four-word sequences*) bersama dengan hitungan jumlah kemunculan tiap-tiap tipe tersebut dalam korpus.

Aplikasi AntConc merupakan aplikasi yang dikembangkan oleh Laurance Anthony, Universitas Waseda, Jepang (Anthony, 2011). Aplikasi

AntConc tergolong aplikasi yang mudah digunakan dan dapat diunduh secara gratis.

METODE PENELITIAN

Penelitian ini merupakan penelitian kualitatif. Metode yang digunakan yaitu metode statistik deskriptif. Metode statistik deskriptif merupakan sekumpulan prosedur dasar yang terdiri atas pengumpulan, pengorganisasian, penyajian, analisis, dan interpretasi data (Mangkuatmodjo, 2003). Statistik deskriptif sangat diperlukan untuk mendeskripsikan data-data yang telah dikumpulkan.

Perkembangan teknologi saat ini memungkinkan analisis data bahasa menggunakan metode statistik deskriptif. Penelitian-penelitian yang berhubungan dengan penyusunan korpus bahasa dapat mengandalkan metode statistik deskriptif untuk membaca data-data bahasa yang telah diolah menggunakan *software/aplikasi* bahasa. Data-data bahasa yang berupa kosakata-kosakata dapat dengan mudah diolah menggunakan aplikasi bahasa, kemudian disesuaikan dengan kebutuhan peneliti.

Penelitian yang dilakukan oleh Citraesmana terhadap penggunaan kata *kita* dan *kami* dalam debat capres juga menggunakan metode statistik deskriptif. Penggunaan metode deskriptif berfungsi untuk menghitung jumlah kemunculan kata *kita* dan *kami* dalam korpus debat capres. Penelitian tersebut juga memaparkan dan mendeskripsikan penggunaan kedua kata tersebut menggunakan aplikasi AntConc (Citraesmana, 2019).

Penelitian ini juga menggunakan aplikasi pembaca korpus, AntConc. Aplikasi tersebut berfungsi untuk menjabarkan isi korpus dengan lebih detail, sehingga kosakata-kosakata yang

ada di dalam korpus dapat dimanfaatkan dalam penyusunan kamus.

Data pada penelitian ini bersumber pada majalah-majalah Fuli yang memuat artikel-artikel berbahasa Melayu Ambon. Artikel-artikel yang dijadikan sumber data yaitu “Sejarah Keberadaan Rumah Tua Matitaputty dalam Tradisi Pelantikan Raja di Negeri Harmalakabesi Amahusu” (Gaspersz, 2014), *Tradisi Makan Samanang* (Loppies, 2015), *Tradisi Cuci Negeri Naku* (Pesiwarissa, 2016), *Asal Mula Mata Air Majapahit di Negeri Ema Huaresi Rehung* (Leimena, 2017), *Beasiswa Studi Sastra, Musik, dan Mereka yang Tidak Disebutkan Namanya* (Johannes, 2017), dan *Lapak Baca: Ruang Baca Ideal Masyarakat* (Pieter, 2018). Artikel-artikel tersebut diketik ulang dalam format MsWord. Setelah selesai diketik, data-data tersebut disimpan dalam format *txt*. Hal ini dilakukan agar data dapat dibaca oleh aplikasi AntConc (Anthony, 2011).

PEMBAHASAN

Aplikasi AntConc pada Majalah Fuli

Penyusunan korpus bahasa Melayu Ambon menggunakan Majalah Fuli akan memanfaatkan aplikasi AntConc. Aplikasi AntConc akan menunjukkan seberapa sering sebuah kata digunakan dalam data yang dimasukkan, kolokasi yang terdapat dalam data, dan daftar perkataan yang digunakan.

Berikut ini akan disajikan langkah-langkah penyusunan korpus bahasa Melayu Ambon menggunakan aplikasi AntConc. Langkah pertama, membuka aplikasi AntConc di laptop. Setelah itu, masukkan file yang akan dijadikan data. Sebelum memasukkan fail ke AntConc (*corpus files*), pastikan file disimpan dalam bentuk *.txt*. Pilih menu *file*, kemudian *open file(s)*, pilih *file* yang

akan dijadikan korpus, dan terakhir pilih *open*. Jika langkah-langkah tersebut dilakukan dengan baik, maka bagian *corpus files* akan berisi data-data yang akan dijadikan korpus. Peneliti dapat melakukan analisis berdasarkan *corpus files* yang telah dimasukkan. Data-data yang telah dimasukkan di AntConc akan terlihat pada gambar 1 berikut ini.

Rank	Freq	Word
1	213	deng
2	205	yang
3	178	pung
4	165	dong
5	154	par
6	129	di
7	126	akang
8	123	xd
9	82	tu
10	79	ada
11	75	orang
12	74	ni
13	60	lagu
14	59	katong
15	53	su
16	51	dia
17	48	itu
18	48	lalu
19	48	negeri
20	47	satu
21	46	seng
22	45	biking
23	45	dar
24	43	bisa

Gambar 1 Hasil Input Data *Word List*

Berdasarkan gambar 1 diketahui bahwa korpus data majalah Fuli memiliki 1.399 tipe kata, dan 7.005 *word tokens* (karakter kata). Data yang disajikan pada gambar 1, diurutkan berdasarkan jumlah terbanyak hingga paling sedikit kemunculannya dalam korpus. Selain berdasarkan jumlah kemunculan kata dalam korpus, pengurutan kata juga dapat dilakukan berdasarkan abjad. Jika ingin mengurutkan kata-kata berdasarkan abjad, maka peneliti dapat memilih pilihan urutan berdasarkan kata (*sort by word*) di bagian kanan bawah aplikasi AntConc.

Jika diamati lebih lanjut, aplikasi AntConc memiliki enam menu pilihan analisis selain *Word List*. Menu-menu lainnya yaitu *Concordance*, *Concordance Plot*, *File View*, *Cluster/N-Grams*, *Collocates*, dan *Keyword List*. Setiap menu-menu tersebut memiliki

fungsi yang berbeda-beda, bergantung pada kebutuhan analisis kata dalam korpus.

Menu *concordance* berfungsi untuk melihat posisi kata dalam kalimat. Peneliti dapat memanfaatkan menu ini untuk menentukan kelas kata sebuah kata dalam korpus yang sedang dibuat. Selain untuk menentukan kelas kata, menu *concordance* juga dapat dimanfaatkan untuk melihat contoh penggunaan kata dalam kalimat. Panjang contoh penggunaan kata dalam menu *concordance* dapat ditentukan melalui menu *Search Window Size*.

Menu *Concordance Plot* berfungsi untuk melihat posisi penyebaran suatu kata dalam teks. Jumlah penyebaran penggunaan kata bergantung pada jumlah munculnya sebuah kata pada korpus data yang dianalisis. Semakin sering sebuah kata muncul di dalam korpus, maka jumlah garis persebaran kata dalam *Concordance Plot* semakin banyak.

Menu ketiga yaitu *File View*. Menu tersebut muncul untuk memudahkan pengguna aplikasi AntConc untuk melihat letak penggunaan sebuah kata dalam badan teks. Pengguna program dapat dengan mudah mengecek penggunaan kata dalam konteks yang lebih luas. Hal tersebut berbeda dengan fitur yang ditawarkan pada menu *concordance*, yang hanya menunjukkan contoh penggunaan kata dalam sebuah kalimat.

Menu berikutnya yaitu *Clusters/N-Grams*. Peneliti dapat memanfaatkan menu tersebut untuk melihat pasangan-pasangan kata yang diikuti dengan jumlah kemunculannya dalam korpus. Menu *Cluster/N-Grams* dapat menunjukkan kecenderungan sebuah kata melekat pada kata lainnya sehingga kita mampu menentukan kelas kata dengan baik.

Menu *Collocate* membantu peneliti untuk mengetahui kombinasi penggunaan kata dalam korpus yang diteliti. Berdasarkan kombinasi kata-kata terjadi berulang-ulang, maka peneliti mengetahui kombinasi kelas kata pada satu kata dengan kata lainnya. Semakin besar korpus yang digunakan, maka semakin banyak contoh-contoh kombinasi kata yang dapat dilihat, sehingga memudahkan peneliti untuk menentukan variasi pembentukan kata.

Menu yang terakhir yaitu *Keyword List*. Menu tersebut dapat membantu peneliti untuk membandingkan kata-kata yang ada di satu korpus dengan korpus acuan lainnya. Agar dapat membandingkan kata dalam *Keyword List*, peneliti harus menyediakan beberapa bentuk korpus yang dapat dioperasikan dalam menu tersebut.

Kosakata-Kosakata pada Korpus

Setelah memasukkan data korpus pada aplikasi AntConc, penulis menemukan kosakata-kosakata yang berhubungan satu dan lainnya. Selain itu, data-data yang telah dimasukkan dapat digunakan sebagai dasar penyusunan kamus. Pekamus dapat memanfaatkan menu *Word List* untuk menunjukkan daftar kata-kata yang dapat dimasukan ke kamus, agar kata-kata yang diinginkan diurutkan berdasarkan abjad (a—z), maka terlebih dahulu harus diurutkan berdasarkan kata (*short by word*).

Tabel 1
Hasil pengurutan kosakata berdasarkan abjad

No	Freq	Lema	No	Freq	Lema
1	1	abad	7	126	akang
2	18	abis	8	1	ale
3	2	adang	9	3	alia
4	19	adat	10	2	alus
5	2	ade	11	6	amper
6	2	ading	12	16	ana

13	19	antua	28	31	jua
14	3	atorang	29	27	kalo
15	1	bagabung	30	4	kapata
16	5	bageang	31	28	kapitan
17	9	bagini	32	18	karja
18	6	bagitu	33	38	kase
19	9	baileu	34	37	la
20	10	bajalang	35	23	laeng
21	2	bakira	36	24	lia
22	17	huaresi	37	15	macang
23	6	iko	38	16	malesi
24	15	ina	39	10	nea
25	1	inini	40	9	paleng
26	3	jalang	41	154	par
27	6	jao	42	1	tabaus

Tabel 1 memuat contoh hasil pengolahan data korpus berdasarkan abjad. Berdasarkan hasil analisis menggunakan AntConc, jumlah lema yang penulis dapatkan dari korpus majalah *Fuli* yaitu 1.399 *word tokens* (karakter kata). Data-data tersebut penulis sortir terlebih dahulu untuk dijadikan referensi penyusunan kamus melayu Ambon.

Kemunculan kosakata-kosakata yang terdapat pada tabel 1 sangat bervariasi. Terdapat kata-kata yang muncul hanya satu kali, tetapi ada juga kata-kata yang sering muncul. Kosakata-kosakata yang muncul hanya satu kali tidak lantas tidak dapat dijadikan data dalam kamus yang akan dibuat.

Salah satu contoh kosakata yang hanya muncul satu kali dalam korpus majalah *Fuli* yaitu kata *ale*. Kata *ale* merupakan kata yang sangat populer di masyarakat Maluku. Meskipun kata tersebut tidak terlalu populer penggunaannya dalam korpus, tetapi kata *ale* harus dimasukan pada referensi kosakata penyusunan kamus bahasa Melayu Ambon. Kata-kata lainnya yang dianggap penting, tetapi hanya muncul satu kali yaitu *tabaus*, *bagabung*, *abad*, dan *inini*.

Setelah mengetahui kosakata-kosakata apa saja yang ada dalam

korpus, peneliti mencoba melihat posisi kata dalam kalimat dan mengetahui kelas katanya. Agar kedua hal tersebut dapat terlaksana, peneliti menggunakan menu *Concordance* pada aplikasi AntConc. Kata-kata yang akan dijadikan contoh yaitu kata *akang*, *neces*, dan *paleng*.

Hit	KWIC	File
1	anang, Carita satu ni akang mulai dar pulau Seram	Data Maja
2	, dong dap alia pulau satu, akang nama pulau Banda, lalu	Data Maja
3	lalu dong singga sadiki di akang pulau Banda ni par lia	Data Maja
4	a ni par lia pulau ini kata akang bisa par dong tinggal	Data Maja
5	pung dusung mar kalapa ini akang paleng tinggi, mar ane	Data Maja
6	saja deng tujuan mau ambel akang pung aer par kasi manc	Data Maja
7	tar buang kalapa, dia kele akang la bawa turung palang-	Data Maja
8	er mamasa tuh, antua giling akang alus-alus la antua gos	Data Maja
9	ng badang-badang. Kuning ni akang par ganti bada. Karna	Data Maja
10	sambilang panggall, kayo ini akang pung nama kayo maluwar	Data Maja

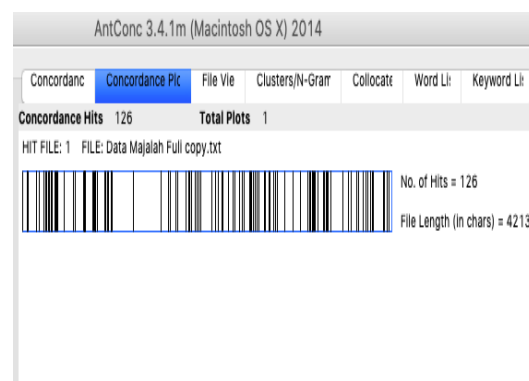
Gambar 2. Hasil analisis *Concordance*

Pada gambar 2 dapat diketahui hasil analisis *Concordance* pada kata *akang*. Kata *akang* merupakan kata yang muncul sebanyak 126 kali sehingga kata tersebut memiliki 126 kali variasi penggunaan kata. Kata *akang* pada contoh kalimat 1 (*Tradisi makan Samanang. Carita satu ni akang mulai dar Pulau Seram yang par waktu itu ...*) mengacu pada *tradisi makan Samanang*. Contoh kalimat lainnya yaitu pada kalimat 9 (*kuning ni akang par ganti bada*). Kata *akang* pada kalimat 9 mengacu pada kata *kuning*.

Contoh-contoh yang telah dibahas menunjukkan bahwa kata *akang* memiliki kelas kata pronomina, khususnya pronomina intertekstual. Pronomina intertekstual merupakan kata ganti yang berfungsi mengganti kelas kata nomina yang terdapat dalam sebuah wacana (Kridalaksana, 2007). Kata *akang* hanya akan muncul pada kalimat-kalimat yang memiliki anteseden.

Persebaran penggunaan kata *akang* dalam korpus majalah *Fuli* dapat dilihat menggunakan menu *Concordance Plot*. Pada menu tersebut dapat diketahui

posisi penggunaan kata *akang* dalam kalimat. Pola persebaran penggunaan kata *akang* terlihat pada gambar 3 berikut ini.



Gambar 3. Hasil analisis *Concordance Plot*

Gambar 3 menunjukkan pola persebaran kata *akang* pada korpus majalah *Fuli*. Kata *akang* memiliki kombinasi penggunaan sebanyak 126 pola. Kata *akang* dapat digunakan pada awal, tengah, dan belakang sebuah kalimat. Hal tersebut sesuai dengan kelas katanya yaitu pronomina. Kata-kata yang berkelas kata pronominal dapat terletak di berbagai posisi, bergantung pada antisedennya. Pola persebaran kata *akang* dapat lebih jelas terlihat jika peneliti memilih menggunakan *plot zoom* (pola penyebaran kata dalam kalimat) di atas empat.

Pelacakan contoh kalimat (utuh) kata *akang* pada korpus majalah *Fuli* dapat memanfaatkan menu *File View* atau langsung mengeklik garis-garis pada hasil analisis pada menu *Concordance Plot* (gambar 3).

Kata-kata yang melekat pada kata *akang* dapat dilihat pada menu *Clusters/N-Grams*. Pada menu tersebut peneliti dapat mengetahui kata-kata yang melekat di sebelah kanan dan kiri kata *akang*. Kata-kata tersebut dapat membantu peneliti untuk menentukan kelas kata. Hasil analisis kata *akang* menggunakan menu *Cluster/N-Grams* terlihat pada gambar 4 berikut ini.

Gambar 4. Hasil analisis Cluster/N-Grams kata *akang*

Concordance				Concordance Plot				File View				Clusters			
Total No. of Cluster Types 63				Total No. of Cluster				Total No. of Cluster Types 79				Total No. of Cluster			
Rank	Freq	Range	Cluster	Rank	Freq	Range	Cluster	Rank	Freq	Range	Cluster	Rank	Freq	Range	Cluster
1	39	1	akang pung	1	11	1	ni akang	1	11	1	ni akang	1	11	1	ni akang
2	5	1	akang paz	2	11	1	yang akang	2	11	1	yang akang	2	11	1	yang akang
3	4	1	akang jadi	3	7	1	tu akang	3	7	1	tu akang	3	7	1	tu akang
4	3	1	akang bawazma	4	4	1	manyanji akang	4	4	1	manyanji akang	4	4	1	manyanji akang
5	3	1	akang dalang	5	3	1	bilang akang	5	3	1	bilang akang	5	3	1	bilang akang
6	3	1	akang deng	6	3	1	di akang	6	3	1	di akang	6	3	1	di akang
7	3	1	akang di	7	3	1	ini akang	7	3	1	ini akang	7	3	1	ini akang
8	3	1	akang jua	8	3	1	lia akang	8	3	1	lia akang	8	3	1	lia akang
9	3	1	akang la	9	3	1	tu, akang	9	3	1	tu, akang	9	3	1	tu, akang
10	2	1	akang ada	10	2	1	bakar akang	10	2	1	bakar akang	10	2	1	bakar akang
11	2	1	akang nama	11	2	1	biking akang	11	2	1	biking akang	11	2	1	biking akang
12	2	1	akang negri	12	2	1	deng akang	12	2	1	deng akang	12	2	1	deng akang
13	2	1	akang paleng	13	2	1	kumpul akang	13	2	1	kumpul akang	13	2	1	kumpul akang
14	2	1	akang su	14	2	1	lalu akang	14	2	1	lalu akang	14	2	1	lalu akang
15	2	1	akang tampa	15	2	1	mangarti akang	15	2	1	mangarti akang	15	2	1	mangarti akang

Gambar 4 menunjukkan jumlah variasi penggunaan kata *akang* yang berada di sebelah kanan lebih banyak dibandingkan di sebelah kiri. Kata-kata yang melekat di sebelah kanan kata *akang* berjumlah 79 variasi, sedangkan kata-kata yang berada di sebelah kiri sebanyak 63 variasi. Kata yang paling sering muncul di sebelah kiri kata *akang* yaitu kata *pung*. Sedangkan kata yang paling sering muncul di sebelah kanan kata *akang* yaitu kata *ni* dan *yang*. Ketiga kata tersebut berkelas kata pronomina.

Kata berikutnya yang akan dibahas yaitu kata *neces*. Meskipun kata *neces* hanya muncul dua kali dalam korpus majalah *Fuli*, kata tersebut merupakan kosakata bahasa Melayu Ambon yang sangat populer. Penggunaan kata *neces* pada korpus Majalah *Fuli* terlihat pada gambar 5.

Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocates	Word List	Key
ni dia su abis mandi mar dia balong neces , lalu dia pung mama ambil kuning deng sore. Skarang ana kacil ni dia su neces mar dia balong ada nama, jadi don						

Gambar 5. Hasil analisis Concordance kata *neces*

Berdasarkan gambar di atas, diketahui bahwa kata *neces* memiliki dua contoh penggunaan. Arti contoh kalimat pertama yaitu *dia sudah selesai mandi tapi dia belum neces, lalu ibunya mengambil kuning*, arti contoh kalimat kedua yaitu *sekarang anak kecil ini dia sudah neces tapi dia belum punya nama*. Berdasarkan kedua contoh kalimat tersebut maka arti kata *neces* dalam bahasa Indonesia yaitu cantik, rapi, sehingga kelas katanya adjektiva. Kata *neces* termasuk adjektiva dasar karena dapat diuji dengan kata sangat, agak, lebih, dan paling.

Penggunaan kata *neces* pada kalimat pertama merupakan bentuk yang umum. Kalimat *neces* pada kalimat pertama berarti cantik. Makna kata cantik muncul dari kalimat penjelas (*lalu ibunya mengambil kuning*). Kata kuning merupakan kata kuncinya. Kuning atau kunyit bagi masyarakat Maluku merupakan salah satu bahan untuk mempercantik diri. Kuning biasanya dijadikan bedak, serta dipakai untuk merawat kulit wajah dari sinar matahari.

Kata *neces* pada kalimat kedua berarti rapi. Makna tersebut dikuatkan dengan kalimat sebelum kata *neces* (*sekarang anak kecil ini dia sudah neces*). Kalimat tersebut menggambarkan keterangan anak kecil yang telah mandi dan bersolek, atau dengan kata lain dia telah rapi.

Concordance	Concordance Plot	File View	Clusters/N-Grams	Collocat
Concordance Hits 2 Total Plots 1 HIT FILE: 1 FILE: Data Majalah Fuli copy.txt No. of Hits = 2 File Length (in chars) = 42138				

Gambar 6. Hasil analisis Concordance Plot kata *neces*

Pada gambar 6 terlihat pola persebaran penggunaan kata *neces* pada kalimat, terdapat dua bentuk plot yang

tergambar pada dua garis vertikal. Kedua garis tersebut menunjukkan penggunaan kata *neces* terletak pada awal kalimat. Meskipun terletak di awal kalimat, kata *neces* harus terlebih dahulu menerangkan suatu benda.

Kata terakhir yang akan dibahas yaitu *paleng*. Pada analisis AntConc, kata *paleng* muncul sebanyak sembilan kali. Kata *paleng* merupakan kata yang populer bagi masyarakat Provinsi Maluku, khususnya Kota Ambon. Contoh kalimat kata *paleng* dapat dilihat pada gambar 7.

Hit	KWIC
1	di orang pung dusung mar kalapa ini akang paleng tinggi, mar ana kacil yang bar lahir pung
2	kata orang ada pancuri kalapa, kama papa ni paleng sayang dia pung ana, dia tar buang kalapa,
3	; ada lai kidung Jemaat, deng satu lai yang paleng batu, Nyanyian GPM, buku sabua yang akang p
4	kerusuhan 1999 v012004, dua lagu yang orang-orang paleng suka menyanyi akang dalang ibadah-ibadah tu
5	-karya sastra, lebe-lebe par oras ini, akang paleng bagus kalo katong tau akang pung khasiat. D
6	susa mangarti keadaan. Sama orang mau jadi yang paleng batul sampe-sampe dong biking hal-hal yang
7	siak kaco tarus tagai sama orang rasa diri paleng batul karna dong pung pikirang yang seng bi
8	aca mulai du Unpatti dari Januari 2018. Ada paleng banyak jenis buku yang lapak baca taru par
9	kase batas sampe Maret setiap taong sa. Penulis paleng berharap deng adanya atorang baru Undang-Un

Gambar 7. Hasil analisis Concordance kata *paleng*

Korpus Melayu Ambon memuat sembilan contoh penggunaan kata *paleng*. Arti contoh kalimat pertama yaitu *kelapa ini sangat tinggi, tetapi anak kecil yang baru lahir punya...*; arti contoh kalimat kedua yaitu *karena papa sangat sayang anaknya...*; arti contoh kalimat ketiga yaitu *ada satu lagi yang paling baru, nyanyian GPM..*; arti kalimat keempat yaitu *orang-orang sangat suka menyanyi...*; dan arti kalimat kelima yaitu *terlebih lagi saat ini, itu sangat bagus jika kita tau khasiatnya....* Berdasarkan terjemahan contoh-contoh kalimat di atas diketahui arti kata *akang* yaitu sangat.

Kata *paleng* memiliki kelas kata adverbial. Adverbial atau disebut juga kata keterangan merupakan kata yang menjelaskan verba, adjektiva, atau adverbial lainnya. Kata *paleng* pada kalimat pertama menjelaskan kata *tinggi* (adj), kalimat kedua menjelaskan kata

sayang (v), kalimat ketiga menjelaskan kata *baru* (adj), kalimat keempat menjelaskan kata *suka* (adj), dan kalimat kelima menjelaskan kata *bagus* (adj).

Concordance Hits: 9, Total Plots: 1
 HIT FILE: 1 FILE: Data Majalah Full copy.txt
 No. of Hits = 9
 File Length (in chars) = 42138

Gambar 8. Hasil analisis Concordance Plot kata *paleng*

Gambar 8 menunjukkan pola persebaran penggunaan kata *paleng* dalam korpus majalah *Fuli*. Kata *paleng* dapat berada di bagian awal, tengah, dan belakang. Hal tersebut sesuai dengan fungsi kelas kata adverbial, sebagai penjelas verba, adjektiva, atau adverbial lainnya.

Kelas kata adverbial pada kata *paleng* juga semakin jelas terlihat pada hasil analisis menggunakan menu *Clusters/N-Grams* (gambar 9). Peneliti dapat mengecek penggunaan kata jika dibandingkan dengan kata di sebelahnya jika menggunakan menu tersebut. Gambar 9 menunjukkan kata-kata yang melekat, baik di sebelah kiri maupun di sebelah kanan, pada kata *paleng*.

Rank	Freq	Range	Cluster
1	1	1	ada paleng
2	2	1	akang paleng
3	1	1	diri paleng
4	1	1	ni paleng
5	1	1	orang paleng
6	1	1	penuis paleng
7	2	1	yang paleng

Rank	Freq	Range	Cluster
1	2	1	paleng batul
2	1	1	paleng bagus
3	1	1	paleng banyak
4	1	1	paleng baru
5	1	1	paleng berharap
6	1	1	paleng sayang
7	1	1	paleng suka
8	1	1	paleng tinggi

Gambar 9. Hasil analisis Clusters/N-Grams kata *paleng*

Pada gambar 9 diketahui terdapat tujuh variasi kata yang melekat di sebelah kanan, dan delapan variasi kata yang melekat di sebelah kiri kata *paleng*. Gambar 9 juga menunjukkan kata *paleng* memiliki kelas kata adverbia.

Analisis-analisis yang telah dilakukan di atas menggambarkan penyusunan korpus bahasa daerah bukanlah hal yang mustahil dilakukan. Meskipun demikian, pemilihan dan penentuan sumber korpus sangat berperan dalam menjaring data-data korpus bahasa daerah yang dibutuhkan.

PENUTUP

Penyusunan korpus bahasa daerah merupakan hal yang tergolong jarang dilakukuan. Salah satu faktor penyebabnya yaitu minimnya sumber tertulis yang biasanya dijadikan sumber data penyusunan korpus. Meskipun demikian, hal tersebut jangan dijadikan penghambat bagi pengembangan dan penyusunan korpus bahasa daerah.

Pemanfaatan aplikasi pembaca korpus sangat membantu peneliti untuk menghasilkan data korpus. Aplikasi AntConc merupakan aplikasi penyusun korpus yang mudah untuk digunakan, serta membantu peneliti untuk menyusun kosakata-kosakata dalam bahasa Melayu Ambon. Menu-menu yang terdapat dalam aplikasi tersebut sangat memudahkan penggunaanya untuk memahami analisis kata-kata yang terdapat dalam korpus. Selain untuk menganalisis kata-kata dalam korpus, menu *Word List* dapat difungsikan sebagai data dasar penyusunan kamus.

Berdasarkan hasil analisis menggunakan aplikasi AntConc diketahui jumlah kata yang terdapat dalam korpus majalah *Fuli* yaitu sebanyak 1.399. Setiap kata-kata tersebut memiliki variasi dalam penggunaannya. Agar memudahkan

pembaca dalam memahami penggunaan aplikasi AntConc, penulis tidak membahas 1.399 kata yang terdapat dalam korpus majalah *Fuli*, tetapi penulis hanya mengambil tiga kosakata. Kosakata-kosakata tersebut yaitu *akang*, *neces*, dan *paleng*. Berdasarkan hasil analisis, diketahui kombinasi kata leksikal dan konstruksi gramatikal dari ketiga kata tersebut.

Pengembangan korpus bahasa daerah harus terus-menerus digalakkan demi terciptanya dokumentasi bahasa daerah. Pemilihan data bahasa daerah lainnya di Provinsi Maluku sebagai data korpus merupakan saran penelitian berikutnya. Selain sumber tertulis sebagai sumber datanya, penelitian berikutnya dapat menysasar data lisan bahasa daerah sebagai sumber data penyusunan korpus bahasa daerah.

DAFTAR PUSTAKA

- Anthony, L. (2011). *AntConc. AntConc (Version 3.2.4) [Computer Software]. Tokyo, Japan: Waseda University. Available from Http://Www.Laurenceanthony.Net/*.
- Citraresmana, E. (2019). Konsep makna kita dan kami dalam debat capres bulan Januari 2019: Kajian semantik kognitif melalui studi korpus (the conceptual meanings of kita and kami in Indonesian). *Metalingua*, 17(2), 103–112. <https://doi.org/http://dx.doi.org/10.26499/metalingua.v17i2.299>
- Collins, J. T. (2018). *Penelitian bahasa di Maluku*. (Asrif, Ed.) (Pertama). Ambon: Kantor Bahasa Maluku.

- Fajri, N. C. (2018). Korpus data bahasa Indonesia dalam teks Indo-Belanda. In E. dkk Suzanti (Ed.), *Leksikografi di Era Digital* (pp. 165—177). Jakarta: Badan Pengembangan dan Pembinaan Bahasa. Retrieved from <http://badanbahasa.kemdikbud.go.id/lamanbahasa/sites/default/files/FInal SLI 2018 Cetak.pdf>
- ForSMAtu. (2016). *Kamus populer bahasa Ambon*. (J. J. Limba, Ed.). Ambon: Forum Alumni SMA N 1 Ambon.
- Gaspersz, M. (2014). Sejarah keberadaan rumah tua Matitaputty dalam tradisi pelantikan raja di Negeri Harmalakabesi Amahusu. *Fuli*, 4—5.
- Hizbullah, N., Fazlurrahman, & Fauziah, F. (2016). Linguistik korpus dalam kajian dan pembelajaran bahasa Arab di Indonesia. *Konferensi Nasional Bahasa Arab (KONASBARA) II*.
- Johannes, W. (2017, November). Beasiswa studi sastra, musik, dan mereka yang tidak disebutkan namanya. *Fuli*, 25—33.
- Jones, R. J. (2008). *Loan words in Indonesian and Malay*. Jakarta: Yayasan Obor Indonesia.
- Kridalaksana, H. (2007). *Kelas kata dalam bahasa Indonesia* (Kelima). Jakarta: PT Gramedia Pustaka Utama.
- Leimena, D. Y. (2017, November). Asal mula mata air Majapahit di Negeri Ema Huaresi Rehung. *Fuli*, pp. 9—14.
- Loppies, A. (2015). Tradisi makan Samanang. *Fuli*, 21.
- Mangkuatmodjo, S. (2003). *Pengantar statistik*. Jakarta: Rineka Cipta.
- McEnery, Tony; Hardie, A. (2012). *Corpus linguistics*. (J. etc Austin, P; Bresnan, Ed.), *Cambridge University Press* (I). United States, New York: Cambridge University Press.
- McEnery, T., & Wilson. (2007). *Corpus linguistics*. Edinburgh: Edinburgh University Press.
- Nurjannah, N. (2018). Keragaman kosakata dalam korpus sastra: Sebuah kontribusi untuk pengembangan leksikografi Indonesia. In E. dkk Suzanti (Ed.), *Leksikografi di Era Digital* (pp. 187—197). Jakarta: Badan Pengembangan dan Pembinaan Bahasa. Retrieved from <http://badanbahasa.kemdikbud.go.id/lamanbahasa/sites/default/files/FInal SLI 2018 Cetak.pdf>
- Pesiwarissa, L. F. (2016, November). Tradisi Cuci Negeri Naku Kecamatan Leitimur Selatan, Kota Ambon. *Fuli*, 16—23.
- Pieter, M. H. E. (2018, November). Lapak baca: Ruang baca ideal masyarakat. *Fuli*, 10—13.
- Sujaini, H. (2018). Peningkatan akurasi penerjemah bahasa daerah dengan optimasi korpus paralel. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, 7(1). <https://doi.org/10.22146/jnteti.v7i1.394>
- Takaria, D., P. (1998). *Kamus bahasa Melayu Ambon-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.

Yuliawati, S. (2018). Perempuan atau wanita? Perbandingan berbasis korpus tentang leksikon berbasis gender. *Paradigma, Jurnal Kajian Budaya*, 8(1), 53. <https://doi.org/10.17510/paradigma.v8i1.227>