

LINGUISTIK KOMPUTASI SEBAGAI DISIPLIN ILMU DAN RESPONS AKADEMIK DI LINGKUNGAN PERGURUAN TINGGI DI INDONESIA

Oleh
Musthofa

Fakultas Adab dan Ilmu Budaya UIN Sunan Kalijaga
Jl. Laksda Adisucipto, Sleman, DI Yogyakarta
Surel: musthofa.bsa@uin-suka.ac.id

Abstract

This paper describes an epistemological perspective of a new discipline in applied linguistics studies called "computational linguistics," the challenge and response faced by Indonesian linguists, and their problems in developing it. The analysis is carried out by utilizing epistemological concepts and theories, combined with Gianluca Carnabuci's theory called "knowledge hybridization" and the "challenge and response" theory proposed by Arnold J. Toynbee. Because this paper focuses on two problems, the data are in the form of two things. First is data relating to computational linguistics based on an epistemological perspective. Second is data relating to the response of universities to the emergence of computational linguistics as a new science. The universities are focused on those who have linguistics study programs, not only those under the Ministry of Religious Affairs but also those under the Ministry of Education and Culture. The results show that epistemologically, computational linguistics has the validity to be called a scientific discipline. However, the response of Indonesian academics to its existence is still inadequate. This is related to the various challenges they face. For this reason, it is necessary to make multiple efforts to overcome it.

Keywords: *computer, computational linguistics, language.*

Abstrak

Tulisan ini akan memaparkan perspektif epistemologis suatu disiplin ilmu baru dalam kajian linguistik terapan yang disebut linguistik komputasi, tantangan dan respons yang dihadapi oleh para linguis Indonesia, serta permasalahan dalam mengembangkannya. Analisis dilakukan dengan memanfaatkan konsep-konsep dan teori epistemologis, yang dipadu dengan teorinya Gianluca Carnabuci yang disebut “hibridasi ilmu pengetahuan” serta teori “tantangan dan respons” dari Arnold J. Toynbee. Karena tulisan ini berfokus pada dua hal, maka data berupa dua hal. Pertama, data yang berupa fakta-fakta terkait linguistik komputasi dari perspektif epistemologi. Kedua, data yang berkaitan dengan respons perguruan tinggi terhadap kemunculan linguistik komputasi sebagai sebuah ilmu baru. Perguruan tinggi di sini difokuskan perguruan tinggi yang memiliki prodi-prodi linguistik, baik perguruan tinggi yang berada di bawah Kementerian Agama maupun di bawah Kementerian Pendidikan dan Kebudayaan. Hasil penelitian menunjukkan bahwa secara epistemologis, linguistik komputasi memiliki keabsahan untuk disebut sebagai sebuah disiplin keilmuan. Namun, respons dunia akademik Indonesia terhadap keberadaannya masih belum memadai. Hal itu terkait dengan berbagai tantangan yang dihadapinya. Untuk itu, perlu dilakukan berbagai upaya untuk mengatasinya.

Kata kunci: komputer, linguistik komputasi, bahasa.

A. PENDAHULUAN

Setiap disiplin ilmu pengetahuan akan senantiasa mengalami pertumbuhan dan perkembangan seiring dengan perkembangan pemikiran manusia. Sebuah disiplin ilmu pengetahuan bisa jadi awalnya cukup sederhana, dan hanya merupakan jawaban atas pertanyaan filosofis “apa, siapa, mengapa, kapan, di mana, dan bagaimana” yang berkaitan dengan sebuah objek tertentu. Namun, karena objek beserta unsur pembangunnya tersebut terus-menerus dipertanyakan dan dicari jawabannya, maka terciptalah sebuah disiplin ilmu pengetahuan yang semakin lama semakin kompleks. Apalagi jika sebuah disiplin ilmu pengetahuan digabungkan dengan disiplin ilmu pengetahuan lain, maka akan menghasilkan sebuah disiplin ilmu pengetahuan baru yang semakin kompleks.

Disiplin ilmu filsafat barangkali pada awalnya adalah murni filsafat. Akan tetapi, dalam perkembangannya disiplin ilmu ini kemudian dikembangkan dan diaplikasikan dalam berbagai disiplin ilmu lain seperti bahasa atau sejarah sehingga menghasilkan disiplin ilmu “filsafat bahasa” atau “filsafat sejarah”. Disiplin ilmu biologi dan kimia pada awalnya memiliki objek kajiannya masing-masing, yakni makhluk hidup dan unsur-unsur yang ada di alam semesta. Namun ketika kedua disiplin ilmu tersebut digabungkan, maka menghasilkan disiplin ilmu baru yang dikenal dengan “biokimia”. Demikian halnya dengan disiplin ilmu “linguistik” dan “komputer”, yang pada dasarnya keduanya memiliki objek kajiannya masing-masing yang berbeda antara satu dengan yang lainnya. Linguistik adalah ilmu tentang bahasa, atau studi tentang bahasa secara ilmiah (Verhaar 2006, 886), sedangkan ilmu komputer adalah ilmu yang membahas mengenai cara kerja mesin yang disebut “komputer” dengan berbagai bentuk rekayasannya, yang bisa diprogram, dan dapat digunakan untuk mengolah, menyimpan, dan mendapatkan kembali, serta memproses data (Webster 1994). Namun, dalam aplikasi dan pengembangannya, ternyata komputer tidak bisa dilepaskan dari bahasa, dan bahasa juga tidak bisa dilepaskan dari komputer. Hal ini karena dalam realitasnya, komputer membutuhkan bahasa dalam memprogram sistem perangkat lunak (*software*)nya di satu sisi, dan di sisi lain, komputer dengan berbagai kecanggihannya ternyata juga dapat diprogram sedemikian rupa sehingga mampu mengolah dan memproses bahasa sebagaimana yang dilakukan oleh manusia sebagai pemilik bahasa (Musthofa 2010). Karenanya, muncul berbagai upaya studi ilmiah dari para ahli yang memadukan antara disiplin linguistik dengan disiplin ilmu komputer. Perpaduan antara disiplin linguistik dan disiplin ilmu komputer di dalam studi bahasa inilah yang kemudian menghasilkan sebuah disiplin baru dalam linguistik yang dikenal dengan “linguistik komputasi (*computational linguistics*)”. Sebuah model pengembangan disiplin ilmu pengetahuan yang oleh Gianluca Carnabuci (2005, 13) disebut sebagai hibridasi atau perpaduan atau pencangkakan ilmu pengetahuan (*knowledge hybridization*)”.

Hadirnya kajian baru dalam bidang linguistik yang disebut “linguistik komputasi” (*computational linguistics*) merupakan sesuatu

yang tak dapat dielakkan dan dihindari. Kajian-kajian linguistik komputasi telah menjadi bagian utama kajian linguistik dan banyak dilakukan oleh para ahli linguistik di negara-negara Barat, sementara di negara kita Indonesia, kajian linguistik komputasi belum mendapat tempat di hati para ahli linguistik. Hal ini paling tidak, tampak pada kurikulum prodi linguistik di berbagai perguruan tinggi di negeri ini yang belum memasukkan kajian linguistik komputasi sebagai bagian dari kurikulum dan kajian linguistik. Beberapa perguruan tinggi yang pernah ditelusuri kurikulumnya oleh penulis, seperti Universitas Gadjah Mada, Universitas Diponegoro, Universitas Atmajaya, Universitas Negeri Malang, Universitas Sam Ratulangi, semuanya belum memasukkan mata kuliah linguistik komputer sebagai bagian dari kurikulumnya. Oleh karenanya, persoalan ini menjadi cukup menarik untuk dikaji dan diteliti.

Berdasarkan uraian di atas, maka masalah yang akan menjadi fokus pembahasan dalam penelitian ini adalah: 1) Apa hakikat keberadaan dan keabsahan linguistik komputasi dan signifikansinya dalam kajian linguistik dari sudut pandang epistemologi? 2) Seperti apa realitas kajian linguistik komputasi di berbagai perguruan tinggi di negeri ini? 3) Apa yang harus dilakukan oleh para ahli linguistik di dalam menyikapi dan merespons kehadiran linguistik komputasi? 4) Problem dan tuntutan apa yang dihadapi oleh ahli linguistik di dalam mengkaji dan menekuni kajian linguistik komputasi?

B. EPISTEMOLOGI LINGUSITIK KOMPUTASI DAN RESPONS PERGURUAN TINGGI DI INDONESIA TERHADAP KEBERADAANNYA

1. Epistemologi dan Perkembangan Ilmu Pengetahuan

Ilmu pengetahuan tidak pernah stagnan. Ia akan senantiasa berkembang sesuai dengan tuntutan zaman. Sebuah disiplin ilmu pengetahuan bisa dikembangkan di dalam disiplin ilmu itu sendiri, tetapi juga bisa dikembangkan dengan cara dipadukan dengan disiplin ilmu lain di luar ilmu tersebut sehingga muncul disiplin ilmu baru sebagaimana disiplin linguistik komputasi. Sebagai ilmu baru, disiplin ilmu ini menarik untuk

dikaji secara epistemologi sehingga kebenaran dan keabsahan terhadap keberadaan linguistik komputasi ini bisa diakui dan menjadi jelas.

Untuk membahas dan menganalisis kebenaran dan keabsahan terhadap keberadaan linguistik komputasi ini, penulis akan menggunakan pendekatan epistemologi (teori ilmu pengetahuan) yang dipadu dengan teorinya Gianluca Carnabuci yang disebut hibridasi ilmu pengetahuan (*knowledge hybridization*) dan teori tantangan dan respons (*challenge and response*) dari Arnold J. Toynbee. Pendekatan epistemologi dan hibridasi ilmu pengetahuan dipakai untuk membedah dan menganalisis keberadaan dan keabsahan disiplin ilmu linguistik komputasi, sedangkan pendekatan tantangan dan respons dipakai untuk membedah mengenai keberadaan disiplin ilmu linguistik komputasi sebagai sebuah tantangan bagi para ahli linguistik dan problem apa yang mesti dihadapi untuk responsnya.

Epistemologi adalah sebuah teori yang berkaitan dengan penyelidikan terhadap sifat dasar dan dasar-dasar ilmu pengetahuan. Pertanyaan mengenai “Apa yang bisa kita ketahui dan bagaimana kita mengetahuinya?” merupakan pertanyaan sentral dan utama di dalam filsafat dan ilmu pengetahuan yang menjadi topik utama di dalam epistemologi. Hal tersebut biasanya berkaitan dengan ide-ide atau gagasan lain yang berupa kepercayaan, pemahaman, alasan-alasan atau pertimbangan atas keberadaannya, penilaian, sensasi, persepsi, institusi, dan pembelajaran berkaitan dengan ilmu pengetahuan tersebut (Lacey 1996, 19). Fokus utama dari epistemologi adalah bentuk-bentuk konsep, teori, dan berbagai problem utama di dalam memahami ilmu pengetahuan. Secara historis, pembenaran dan alasan pembenaran, yang biasanya berkaitan dengan alasan untuk memercayai ilmu pengetahuan, data-data, dan jaminan atas kebenarannya, merupakan sesuatu yang sangat penting dalam epistemologi (Audi 1999, xi). Dalam konsep dan kerangka pemikiran inilah, hakikat keberadaan dan keabsahan disiplin linguistik komputasi akan dibahas dan dianalisis.

Di samping itu, untuk melengkapi kerangka epistemologis di atas, penulis juga akan menggunakan kerangka pemikiran Gianluca Carnabuci, yaitu hibridasi ilmu pengetahuan. Dalam pemikirannya, dinamika perkembangan ilmu pengetahuan tidak bisa dilepaskan dari

jaringan aktual pola-pola perpaduan atau pencangkakan dengan atau antara berbagai domain ilmu pengetahuan (Carnabuci 2005, 13). Pendapat ini didasarkan pada pernyataan Abbott Payson Usher (Carnabuci 2005, 11), bahwa pemikiran-pemikiran dan penciptaan hal baru dalam ilmu pengetahuan menemukan bentuknya yang berbeda karena adanya asimilasi konstruktif dari elemen-elemen ilmu pengetahuan yang telah ada, ke dalam suatu sintesis baru, pola baru, dan konfigurasi baru ilmu pengetahuan.

Di sisi lain, hadirnya disiplin ilmu pengetahuan baru, yang hal ini merupakan bagian dari tuntutan perkembangan ilmu pengetahuan, akan senantiasa menjadi tantangan bagi para ilmuwan. Apa yang harus mereka lakukan dan bagaimana mereka merespons atau menolak hadirnya disiplin ilmu pengetahuan baru tersebut adalah sesuatu yang lumrah terjadi di dunia ilmu pengetahuan. Berkaitan dengan hal ini, Arnold J. Toynbee mengemukakan sebuah teori yang dikenal dengan tantangan dan respons (*challenge and response*). Ia berpendapat bahwa hadirnya budaya baru (dalam hal ini hadirnya disiplin ilmu pengetahuan baru) merupakan tantangan yang bisa menjadi ancaman bagi eksistensi budaya yang ada (ilmu pengetahuan) jika hal itu tidak bisa dihadapi dan direspons secara kreatif. Namun, jika budaya baru tersebut bisa direspons secara kreatif, maka hal ini akan membawa dampak yang baik bagi perkembangan dan kemajuan sebuah budaya, (Toynbee 1972, 97; dan Burke 1992, 27). Karena pada dasarnya, manusia adalah makhluk budaya, maka dengan kemampuan dan daya pikirnya, ia akan senantiasa mempelajari simbol-simbol dan menerima informasi-informasi baru yang ia temukan, kemudian diadaptasikan dengan dirinya dan kelompoknya sesuai dengan kebutuhannya (Ritzer 1996, 210). Dengan menggunakan kerangka pemikiran ini, maka hadirnya sebuah disiplin ilmu pengetahuan baru yang disebut linguistik komputasi sebagai sebuah tantangan bagi para ahli linguistik, apa yang harus mereka lakukan, bagaimana mereka merespons atau menolak, dan problem apa yang dihadapi terhadap hadirnya disiplin ilmu pengetahuan tersebut dapat dielaborasi dan dibahas.

2. Epistemologi Linguistik Komputasi

Epistemologi merupakan cabang dari filsafat yang membahas tentang asal, sifat, karakter, dan jenis pengetahuan. Ada beberapa pertanyaan sentral berkaitan dengan epistemologi pengetahuan, yang di antaranya adalah: 1) Apa ilmu pengetahuan itu? yang hal ini berkaitan dengan definisi dan konsep ilmu pengetahuan, 2) Apa tingkat dari ilmu pengetahuan? yang hal ini berkaitan dengan apa sebenarnya macam dari diadakannya ilmu pengetahuan, 3) Apa sumber-sumber pengetahuan? yang hal ini berkaitan dengan bagaimana pengetahuan itu didapat, dan 4) Apa sejatinya ilmu pengetahuan itu? (Routledge 2000, 250)

Berkaitan dengan pertanyaan epistemologis pertama yang berkaitan dengan pengertian linguistik komputasi, Igor A. Bolshakov dan Alexander Gelbukh (2004, 25) menyatakan bahwa:

Computational linguistics might be considered as a synonym of automatic processing of natural language, since the main task of computational linguistics is just the construction of computer programs to process words and texts in natural language.

Pendapat yang lain dikemukakan oleh Ruslan Mitkov. Ia menyatakan bahwa “*computational linguistics is an interdisciplinary field concerned with the processing of language by computer*” (Mitkov 2004, ix). Sedangkan dalam kamus *Routledge Dictionary of Language and Linguistics*, linguistik komputasi didefinisikan sebagai:

“Discipline straddling linguistics and (applied) computer science that is concerned with the computer processing of natural languages on all levels of linguistic description” (Bussmann 1996, 224).

Berdasarkan beberapa pengertian di atas, maka linguistik komputasi dapat dipahami sebagai sebuah studi ilmiah interdisipliner mengenai bahasa dari sudut pandang komputer, yang menitikberatkan pada rekayasa program aplikasi komputer sehingga komputer mampu melakukan pemrosesan dan pengelolaan terhadap bahasa natural sebagaimana yang biasa dilakukan oleh manusia.

Secara historis, istilah *computational linguistics* (linguistik komputasi) baru muncul pada tahun 1974 sebagai padanan bagi istilah *machine translation* yang dimuat di jurnal *American Journal of*

Computational Linguistics. Meskipun demikian, embrio munculnya disiplin linguistik komputasi telah dimulai sejak tahun 1949 ketika Warren Weaver membuat pernyataan yang sangat terkenal dan mengejutkan banyak kalangan bahwa “*translation by machine might be possible*”. Tahun 1955, setelah melalui berbagai penelitian yang dilakukakannya, ia menegaskan kembali bahwa “*computers might be useful for the solution of world-wide translation problems*”. Kemajuan dari terjemahan mesin ke Linguistik Komputasi terjadi pada tahun 1974 ketika terjemahan mesin dan linguistik komputasi digantikan oleh *American Journal of Computational Linguistics*, yang awalnya hanya muncul dalam bentuk *microfiche*. Kemudian, pada tahun 1980, istilah “*computational linguistics*” mulai dikenal dan eksis sampai sekarang (Mitkov 2004, 25; dan Grishman 1994, 1) Pergeseran besar dalam hampir semua aspek kajian pemrosesan bahasa alami yang dikenal dengan “*computational linguistics*” ini dimulai pada akhir 1980-an dan hampir selesai pada akhir 1995. Hal ini ditandai dengan munculnya isu khusus linguistik komputasi mulai triwulanan pada tahun 1993 (Schubert 2016). Dengan demikian, dapat dipahami bahwa istilah “*computational linguistics*” (linguistik komputasi) baru muncul setelah munculnya istilah “*machine translation*” atau “*mechanical translation*” (penerjemahan oleh mesin atau mekanik).

Di samping itu, secara epistemologis, keberadaan disiplin sebuah ilmu pengetahuan tidak bisa dilepaskan dari sumber-sumber bagi keberadaannya. Menurut Robert Audi (1999, 128), epistemologi, tidak bisa dilepaskan dari sumber-sumber yang membangunnya. Berkaitan dengan ini, ia menyatakan bahwa “*perception, memory, consciousness, and reason are our primary individual sources of knowledge and justification*”. Sedangkan Paul K. Moser (2002, 72) menyatakan bahwa:

“...any sources of knowledge deserve to be called the classical basic sources, the best candidates are perception, memory, consciousness (sometimes called introspection), and reason (sometimes called intuition). Some writers have shortened the list under the heading, “experience and reason”.

Beberapa penulis meringkas daftar sumber tersebut hanya menjadi “pengalaman dan penalaran”.

Munculnya kajian baru dalam bidang linguistik yang disebut linguistik komputasi ini tidak bisa dilepaskan dari hadirnya teknologi komputer dengan berbagai kemampuan dan kecanggihannya dalam melakukan apa saja yang dikehendaki manusia. Hal inilah yang kemudian telah menginspirasi para ahli komputer dan bahasa untuk menjadikan komputer seperti manusia yang mampu melakukan berbagai hal berkaitan dengan bahasa, baik dalam bentuk tulisan maupun ujaran. Para ahli komputer dan bahasa terinspirasi oleh sistem dan cara kerja otak manusia, dan mereka ingin meniru kecerdasan otak manusia tersebut, kemudian menerapkannya pada komputer sehingga komputer menjadi semacam mesin yang cerdas (*smart machine*), yang mampu memproses, mengolah dan menyajikan data kebahasaan secara otomatis, sebagaimana yang biasa dilakukan oleh manusia. Kemampuan komputer di dalam memproses bahasa natural (*natural language*) (Walter 2008), sebagaimana yang dilakukan oleh manusia ini kemudian dikenal dengan "*natural language processing* (pemrosesan terhadap bahasa natural)" yang biasa disingkat menjadi NLP (Richards and Schmidt 2002, 352). Hal inilah yang di antaranya menjadi alasan dan telah mendorong terhadap munculnya disiplin linguistik komputasi.

Dilihat dari sudut pandang asal usul keberadaannya, munculnya linguistik komputasi dilatarbelakangi oleh perkembangan teknologi komputer, yang pada prinsipnya memadukan antara matematika, elektronika, dan bahasa. Basis pengembangan ilmu ini didasarkan pada persepsi dan model matematisasi bahasa yang dikemukakan oleh Chomsky dan model mesin Turing yang dikemukakan oleh Alan Turing, serta didukung oleh pemikiran Warren Weaver mengenai kemampuan mesin komputer di dalam mengolah dan menerjemahkan bahasa manusia secara otomatis (Mitkov 2004, xvii). Mesin komputer yang bersifat elektronik, dalam arti berupa rangkaian sirkuit elektronika, hanya bisa bekerja atas dasar matematika biner dan hanya mengenal bilangan biner "1" dan "0". Kedua angka ini kemudian disebut "*binary digit*" disingkat "*bit*". Apapun yang tampak di layar monitor, baik itu berupa file teks, gambar, suara, video, dan lainnya adalah hasil konversi dan simulasi dari bilangan biner "1" dan "0". Oleh karenanya, satuan yang digunakan

untuk ukuran atau besaran jenis-jenis file tersebut adalah “byte” singkatan dari “binary term” (bentuk biner).

Di samping itu, mesin komputer hanya bisa bekerja di bawah kontrol sistem operasi yang dibangun melalui perintah-perintah (*commands*) berupa sintaks-sintaks (*syntax*) bahasa natural, yang kemudian membentuk sistem operasi. Dengan demikian, maka dapat dikatakan bahwa komputer tidak bisa dilepaskan dari bahasa. Hal ini sesuai dengan pendapat Arto Salomaa (1985, 5) yang menyatakan:

“The basic theory of computation, as well as of its various branches, such as cryptography, is inseparably connected with language theory. The input and output sets of a computational device can be viewed as languages, and - more profoundly - models of computation can be identified with classes of language specifications, in a sense to be made more precise. Thus, for instance, Turing machines can be identified with phrase-structure grammars and finite automata with regular grammars”.

Hubungan erat keterkaitan antara komputer dan bahasa dapat dilihat dengan adanya perintah-perintah kebahasaan di layar monitor seperti “start”, “open”, “close”, “save” “copy” dan sebagainya, yang semua itu menunjukkan bahwa bahasa memang menjadi bagian penting di dalam rancang bangun komputer. Oleh karenanya, di dalam teknologi komputer dikenal adanya pemrograman bahasa (*programming language*) (Microsoft 2002). Melalui pemrograman bahasa yang dipadu dengan logika matematika biner inilah, komputer mampu mengolah berbagai data, baik teks (*text*), suara (*audio*), gambar (*picture*), maupun gambar bergerak (*video*). Hal penting lain dari kecanggihan komputer adalah kemampuannya dalam memproses dan mengolah bahasa natural sebagaimana yang dilakukan oleh manusia sebagai pemilik bahasa.

Komputer, dengan kemampuannya tersebut, seakan memiliki kecerdasan sebagaimana manusia. Pemodelan ini kemudian menghasilkan sebuah kajian dalam bidang ilmu komputer yang dikenal dengan “kecerdasan buatan (*artificial intelegent*)” (Microsoft 2002), yang di antara bagian kajiannya adalah pemrosesan bahasa natural. Atas dasar hubungan erat antara bahasa dan komputer, serta kemampuan komputer di dalam memproses bahasa natural, maka lahir sebuah disiplin baru dalam kajian linguistik yang merupakan hasil hibridasi ilmu pengetahuan, sebuah perpaduan atau pencangkokan ilmu pengetahuan

antara ilmu linguistik dan ilmu komputer. Dari sinilah kemudian, lahir sebuah disiplin ilmu pengetahuan yang disebut linguistik komputasi (*computational linguistics*). Di dalam linguistik umum (*general linguistics*), posisi kajian linguistik komputasi berada dalam kelompok linguistik terapan (*applied linguistics*). Gambaran lebih lanjut mengenai posisi kajian Linguistik Komputasi ini bisa dilihat di artikel berjudul “Linguistik Komputer (*Computational Linguistics*): Model Baru Kajian Linguistik dalam Perspektif Komputer” (Musthofa 2010, 259).

Sebagai sebuah disiplin ilmu pengetahuan, linguistik komputasi memiliki ruang lingkup atau bidang kajian yang diorganisasikan dari hasil eksperimen panjang berdasarkan fakta-fakta dan sumber-sumber pengetahuan, antara ilmu linguistik dan ilmu komputer, sehingga membentuk sebuah bangunan ilmu pengetahuan yang sistematis. Ruang lingkup dan bidang kajian linguistik komputasi mencakup: fonologi, morfologi, leksikon, leksikografi, sintaksis, semantik, pragmatik, pengolahan teks atau pengeditan teks, pemisahan kata secara otomatis, pengecekan ejaan, pengecekan tata bahasa, pengecekan gaya bahasa, dan preferensian antarkata, penggabungan kata, hubungan antar makna kata, pemerolehan informasi, penerjemahan oleh mesin, antar-muka bahasa natural, ekstraksi data faktual, pemahaman terhadap bahasa natural, pengenalan karakter huruf oleh lensa optik, pengenalan ujaran, wacana, dialog, segmentasi teks, peringkasan teks, pelabelan bagian-bagian ujaran, pemenggalan kalimat, pengenalan ujaran, pemaduan pengucapan teks, teknologi stata hingga, resolusi anafora, pembelajaran melalui mesin, bank data bahasa, dan penjawaban atas pertanyaan (Hausser 2001, 21, Bolshakov dan Alexander 2004, 8; dan Mitkov 2004, v-x). Semua ini dikaji dalam perspektif komputer.

Jika memperhatikan berbagai uraian di atas, maka keberadaan kajian linguistik komputasi sudah sangat absah disebut sebagai disiplin ilmu pengetahuan. Karena secara epistemologis, dilihat dari sisi asal-usul keberadaannya, alasan diadakannya, sumber-sumber diadakannya, definisi dan pengertiannya, ruang lingkup kajiannya, sifat dan karakternya, jenis kajiannya, serta kebenaran dan alasan keberadaannya, semuanya sangat jelas tergambar dan memenuhi unsur untuk disebut sebagai ilmu pengetahuan.

3. Keberadaan Kajian Linguistik Komputasi di Berbagai Perguruan Tinggi di Indonesia.

Secara epistemologis, keberadaan linguistik komputasi sah disebut sebagai ilmu pengetahuan. Di dunia Barat ilmu ini telah menjadi isu mutakhir dan telah menjadi bagian penting dalam kajian-kajian akademik di berbagai perguruan tinggi. Namun, di berbagai perguruan tinggi di Indonesia ilmu ini masih menjadi sebuah disiplin ilmu yang langka. Padahal jika melihat data jumlah perguruan tinggi dan program studi (prodi) di Indonesia, maka jumlahnya cukup banyak, dan angkanya bisa dilihat pada tabel di bawah ini.

Tabel 1

Jumlah Perguruan Tinggi (PT) di Indonesia

No	Tahun	Jumlah PT	Jumlah Prodi
1	2011	2.741	12.910***)
2	2016	4.571	18.796***)
3	2021	4.593	29.413*) 28.591**) 26.499***)

*) (Setditjen Dikti 2020, xxii)

***) (“PDDikti - Pangkalan Data Pendidikan Tinggi” n.d.)

*) (Setditjen Dikti 2020, xxii)

Dari sekian banyak prodi yang ada di perguruan tinggi tersebut, prodi Bahasa, Linguistik, Bahasa Arab, Bahasa dan Sastra Arab, dalam berbagai bentuknya, mengalami perkembangan dan peningkatan jumlah yang cukup signifikan sejak tahun 2011 hingga tahun 2021. Gambaran mengenai hal ini bisa dilihat pada tabel berikut.

Tabel 2

Jumlah Prodi Bahasa dan Linguistik di Perguruan Tinggi di Indonesia

No	Nama Prodi	Jumlah Persentase		Jumlah Persentase		Jumlah Persentase	
		Tahun 2011		Tahun 2016		Tahun 2021	
1	Bahasa	613	85,61%	808	82,11%	1033	79,40%
2	Linguistik	15	2,09%	25	2,54%	28	2,15%
3	Bahasa Arab (Pendidikan)	75	10,47%	133	13,52%	207	15,91%
4	Bahasa dan Sastra Arab	13	1,82%	18	1,83%	33	2,54%
	Jumlah	716	100%	984	100%	1301	100%

*Data diambil dari (BAN-PT n.d.)

Berdasarkan tabel di atas, dapat diketahui bahwa di Indonesia, ada sejumlah prodi bahasa dalam berbagai nama dan bentuknya, baik Bahasa, Linguistik, Bahasa Arab (Pendidikan), dan Bahasa dan Sastra Arab (non-pendidikan). Pada tahun 2011 sebanyak 716 prodi, tahun 2016 sebanyak 984 prodi, dan pada tahun 2021 meningkat menjadi 1301 prodi. Lalu, bagaimana keberadaan kajian linguistik komputasi di berbagai prodi pada berbagai perguruan tinggi tersebut? Berkaitan dengan ini, penulis mengambil sampel 19 prodi Bahasa dan Sastra Arab, Bahasa dan Sastra Inggris, dan Bahasa Inggris di berbagai Perguruan Tinggi Keagamaan Islam Negeri (PTKIN) di bawah naungan Kementerian Agama (Kemenag), dan 21 prodi linguistik di berbagai Perguruan Tinggi Negeri (PTN) di bawah naungan Kementerian Pendidikan dan Kebudayaan (Kemendikbud), mulai tahun 2011 sampai tahun 2021. Pada tahun 2011, 19 prodi yang ada di PTKIN yang berada di bawah naungan Kemenag, semuanya masih berada di jenjang S1. Sedangkan 21 prodi linguistik di berbagai PTN di bawah naungan Kemendikbud berada pada jenjang S2 dan S3.

Berdasarkan penelusuran penulis terhadap kurikulum 19 prodi yang ada di PTKIN yang berada di bawah naungan Kemenag pada tahun

2011 sampai tahun 2021, semuanya belum memasukkan kajian linguistik komputasi sebagai mata kuliah prodi. Sedangkan penelusuran yang sama yang dilakukan oleh penulis terhadap kurikulum 21 prodi linguistik di berbagai PTN di bawah naungan Kemendikbud sejak tahun 2011 sampai tahun 2021, yang semuanya berada pada jenjang S2 dan S3, ada beberapa prodi yang pada tahun 2011 sudah memasukkan linguistik komputasi sebagai mata kuliah prodi, namun jumlahnya masih sangat sedikit. Pada tahun 2021, jumlah ini mengalami peningkatan yang cukup signifikan. Gambaran mengenai hal ini bisa dilihat pada tabel berikut.

Tabel 3
Prodi Linguistik dan Kajian Linguistik Komputasi

No	Nama Perguruan Tinggi	Nama Prodi	Jenjang	Ada/Tidak ada	
				2011	2021
1	Universitas Gadjah Mada	Ilmu Linguistik	S-2	Tidak ada	Ada
2	Universitas Indonesia	Ilmu Linguistik	S-3	Tidak ada	Ada
3	Universitas Indonesia	Ilmu Linguistik	S-2	Ada	Ada
4	Universitas Sumatera Utara	Ilmu Linguistik	S-3	Ada	Ada
5	Universitas Sumatera Utara	Ilmu Linguistik	S-2	Tidak ada	Tidak ada
6	Universitas Hasanudin	Ilmu Linguistik	S-3	Tidak ada	Tidak ada
7	Universitas Hasanudin	Ilmu Linguistik	S-2	Ada	Ada
8	Universitas Andalas	Ilmu Linguistik	S-2	Tidak ada	Tidak ada
9	Universitas Diponegoro	Ilmu Linguistik	S-2	Tidak ada	Tidak ada
10	Universitas Sam Ratulangi	Ilmu Linguistik	S-3	Tidak ada	Tidak ada

11	Universitas Sam Ratulangi	Ilmu Linguistik	S-2	Tidak ada	Ada
12	Universitas Udayana	Ilmu Linguistik	S-3	Tidak ada	Ada
13	Universitas Udayana	Ilmu Linguistik	S-2	Tidak ada	Ada
14	Universitas Nusa Cendana	Ilmu Linguistik	S-2	Tidak ada	Tidak ada
15	Universitas Sebelas Maret	Ilmu Linguistik	S-3	Tidak ada	Ada
16	Universitas Sebelas Maret	Ilmu Linguistik	S-2	Tidak ada	Tidak ada
17	Universitas Negeri Jakarta	Linguistik Terapan	S-2	Tidak ada	Ada
18	Universitas Negeri Yogyakarta	Linguistik Terapan	S-2	Tidak ada	Ada
19	Universitas Negeri Medan	Linguistik Terapan Bahasa Inggris	S-2	Tidak ada	Tidak ada
20	Universitas Katolik Indonesia Atma Jaya	Linguistik Terapan Bahasa Inggris	S-3	Tidak ada	Tidak ada
21	Universitas Katolik Indonesia Atma Jaya	Linguistik Terapan Bahasa Inggris	S-2	Tidak ada	Tidak ada

***Data diolah dari berbagai kurikulum yang ada di perguruan tinggi di tabel.**

Berdasarkan data pada tabel di atas, dari 21 prodi linguistik yang ada di berbagai perguruan tinggi, yang semuanya berada dalam jenjang S-2 atau S-3, baru ada tiga prodi yang menawarkan kajian linguistik komputasi, yakni: 1) Prodi Ilmu Linguistik Universitas Indonesia (S-2), 2) Prodi Ilmu Linguistik Universitas Sumatera Utara (S-3), dan 3) Prodi

Ilmu Linguistik Universitas Hasanudin (S-2). Sedangkan pada tahun 2021, sudah ada delapan prodi lagi yang menawarkan kajian linguistik komputasi di dalam kurikulumnya, sehingga jumlahnya menjadi 11 prodi. Prodi-prodi tersebut adalah: 1) Prodi Ilmu Linguistik (S-2) Universitas Gadjah Mada, 2) Prodi Ilmu Linguistik (S-3) Universitas Indonesia, 3) Prodi Ilmu Linguistik (S-2) Universitas Sam Ratulangi, 4) Prodi Ilmu Linguistik (S-3) Universitas Udayana, 5) Prodi Ilmu Linguistik (S-2) Universitas Udayana, 6) Prodi Ilmu Linguistik (S-3) Universitas Sebelas Maret, 7) Prodi Linguistik Terapan (S-2) Universitas Negeri Jakarta, dan 8) Prodi Linguistik Terapan (S-2) Universitas Negeri Yogyakarta. Hanya saja, tidak semua prodi ini menyebut nama mata kuliahnya dengan nama linguistik komputasi; ada 3 prodi yang mata kuliahnya berupa “*machine translation*”, dan “*korpus linguistik*”. Namun demikian, mata kuliah ini masih masuk dalam kajian linguistik komputasi. Dalam bentuk persentase, prodi-prodi yang menawarkan kajian linguistik komputasi dapat dilihat pada tabel berikut.

Tabel 4
Prodi Linguistik dan Persentase Kajian Linguistik Komputasi

No	Keberadaan	2011		2021	
		Jumlah	Persentase	Jumlah	Persentase
1	Ada	3	14,29%	11	52,38%
2	Tidak ada	18	85,71%	10	47,62%
	Jumlah	21	100,00%	21	100,00%

Dari tabel di atas, dapat dipahami bahwa respons prodi-prodi di berbagai perguruan tinggi terhadap kajian linguistik komputasi cukup bagus dan meningkat. Prodi-prodi di perguruan tinggi yang menawarkan kajian linguistik komputasi di dalam kurikulumnya naik dari 14,29% menjadi 52,38% dalam waktu 10 tahun terakhir, atau naik sebesar 38,10%. Hal ini menunjukkan bahwa kajian linguistik komputasi dianggap sebagai sebuah kajian yang cukup penting di prodi-prodi

linguistik. Namun, jika memperhatikan data pada tabel di atas, maka tampak bahwa masih ada 10 prodi, atau 47,62% prodi yang belum merespons keberadaan kajian linguistik sebagai bagian penting. Apalagi jika melihat data kurikulum prodi-prodi yang ada di PTKIN yang berada di bawah naungan Kemenag, belum ada prodi yang merespons pentingnya kajian linguistik sebagai bagian penting dalam kajian linguistik.

4. Kajian Linguistik Komputasi di Berbagai Perguruan Tinggi di Indonesia: Tantangan, Respons, dan Problem Basis Keilmuan

Ilmu pengetahuan baru adalah tantangan baru. Sesuatu yang baru biasanya memunculkan problem dan butuh waktu di dalam memahaminya. Kebaruan linguistik komputasi ini adalah salah satu faktor kenapa kajian ini belum dikembangkan di perguruan tinggi di negeri ini. Untuk meresponsnya dibutuhkan pemikiran lebih dan cara baru yang mungkin belum pernah ada dalam diri seseorang. Problem yang dihadapi oleh kalangan akademisi linguistik di dalam memahami dan mengembangkannya juga tidak ringan. Karenanya, respons terhadap kehadiran linguistik komputasi oleh berbagai perguruan tinggi pun berbeda-beda, dan butuh waktu yang relatif lama dalam pengembangannya. Hal ini sesuai dengan pendapat Arnold J. Toynbee (1972, 97) yang menyatakan bahwa:

“Challenge and response resembles cause and effect only in standing for a sequence of events. The character of the sequence is not the same. Unlike the effect of a cause, the response to a challenge is not predetermined, is not necessarily uniform in all cases, and is therefore intrinsically unpredictable”.

Kecepatan di dalam memperoleh informasi dan juga faktor kesiapan kompetensi diri adalah sebagian faktor yang menyebabkan mengapa kajian linguistik komputasi belum dianggap sebagai sebuah tantangan sehingga belum begitu direspons oleh kalangan akademisi linguistik. Karena untuk menghadapi tantangan dibutuhkan energi, kesiapan mental, dan kompetensi diri yang cukup. Oleh karenanya, respons terhadap tantangan hadirnya ilmu pengetahuan baru, yang dalam hal ini adalah linguistik komputasi, akan menjadi semacam tes atau uji

kompetensi atas kemampuan seorang akademisi di dalam memahami dan mengembangkannya. Ini sesuai dengan pengertian tantangan yang ada di dalam kamus *Cambridge Advanced Learner's Dictionary*. Disebutkan bahwa “*challenge is the situation of being faced with something needing great mental or physical effort in order to be done successfully and which therefore tests a person's ability* (Walter 2008).

Belum diresponsnya kajian linguistik komputasi di kalangan akademisi linguistik di berbagai perguruan tinggi ini sangat mungkin berkaitan dengan problem basis keilmuan yang mereka miliki di dalam memahami dan mengembangkan disiplin kajian linguistik komputasi. Linguistik komputasi adalah sebuah disiplin ilmu pengetahuan baru dalam bidang linguistik yang dikembangkan atas dasar hibridasi ilmu pengetahuan yang menggabungkan antara ilmu linguistik dan ilmu komputer. Untuk memahami dan mengembangkan linguistik komputasi, yang dibutuhkan tidak sekadar teori dan konsep dalam ilmu linguistik, tetapi juga teori dan konsep dalam ilmu komputer. Oleh karenanya, problem utama dan mendasar yang dihadapi oleh akademisi linguistik di dalam merespons, memahami, dan mengembangkan disiplin linguistik komputasi adalah tuntutan kompetensi dan keahlian dalam bidang ilmu komputer di samping kompetensi dan keahlian dalam bidang ilmu linguistik yang sama-sama harus dimiliki oleh akademisi linguistik. Tanpa kompetensi dan keahlian dalam bidang ilmu komputer, sulit rasanya bagi kalangan akademisi linguistik untuk bisa merespons, memahami, dan mengembangkan disiplin linguistik komputasi.

Jika memperhatikan problem dan kebutuhan di atas, maka memang tidak mudah dan tidak ringan tantangan dan problem yang dihadapi oleh akademisi linguistik. Model pembelajaran yang ada di berbagai perguruan tinggi di Indonesia hingga saat ini masih belum memadukan antara dua kompetensi dan keahlian, yakni kompetensi dan keahlian dalam bidang linguistik dan komputer. Kompetensi dan keahlian dalam bidang linguistik memang sudah ada dan dimiliki oleh akademisi di prodi linguistik atau prodi bahasa, sedangkan kompetensi dan keahlian dalam bidang ilmu komputer, umumnya ada pada para akademisi di prodi teknik informatika atau prodi ilmu komputer.

Karenanya, akademisi linguistik memiliki problem basis keilmuan berkaitan dengan kompetensi dan keahlian dalam bidang komputer.

Untuk mengatasi problem tersebut perlu upaya dari institusi atau prodi linguistik untuk mendorong para akademisi linguistik mempelajari linguistik komputasi. Hal ini bisa dilakukan melalui pelatihan dalam bidang linguistik. Di samping itu, untuk mentransformasikan dan mengembangkan disiplin linguistik komputasi, akademisi linguistik haruslah didorong untuk membekali diri dengan mempelajari teori komputasi dan berbagai pemrogramannya, lalu diaplikasikan dalam kajian linguistik komputasi. Kemudian hal ini bisa ditindaklanjuti dengan memasukkan materi teori komputer atau mata kuliah linguistik komputasi sebagai bagian dari muatan kurikulum prodi guna membekali lulusannya. Upaya ini akan dapat dan mampu memenuhi kebutuhan kalangan akademisi perguruan tinggi dalam rangka menjawab tantangan, merespons, dan mengembangkan linguistik komputasi, di satu sisi, dan dapat membekali para lulusannya dengan kompetensi linguistik komputasi, sehingga ke depan ilmu ini akan bisa berkembang. Namun, hal ini tidaklah mudah, dan perlu proses. Kita tunggu saja perkembangannya.

C. SIMPULAN

Hadirnya kajian baru dalam bidang linguistik yang disebut linguistik komputasi (*computational linguistics*) merupakan sesuatu yang tak dapat dielakkan dan dihindari. Kajian-kajian linguistik komputasi telah menjadi bagian utama kajian linguistik dan banyak dilakukan oleh para ahli linguistik di negara-negara Barat, sementara di negara kita Indonesia, kajian linguistik komputasi belum mendapat tempat di hati para ahli linguistik.

Secara historis, istilah *computational linguistics* baru muncul pada tahun 1974 sebagai padanan bagi istilah “*machine translation*”. Meski demikian, embrionya sudah muncul sejak tahun 1949 ketika Warren Weaver membuat pernyataan yang sangat terkenal dan mengejutkan banyak kalangan bahwa “*translation by machine might be possible*”. Pada tahun 1955, ia menegaskan kembali bahwa “*computers might be useful for the solution of world-wide translation problems*”. Kemudian,

pada tahun 1980, istilah *computational linguistics* mulai dikenal dan eksis sampai sekarang. Pergeseran besar dalam hampir semua aspek kajian pemrosesan bahasa alami yang dikenal dengan linguistik komputasi ini dimulai pada akhir 1980-an dan hampir selesai pada akhir 1995.

Disiplin linguistik komputasi dikembangkan atas dasar hubungan erat antara bahasa dan komputer, serta kemampuan komputer di dalam memproses bahasa natural. Maka lahirlah sebuah disiplin baru dalam kajian linguistik yang merupakan hasil hibridasi ilmu pengetahuan, yakni sebuah perpaduan atau pencangkokan ilmu pengetahuan antara ilmu linguistik dan ilmu komputer.

Linguistik komputasi dapat dipahami sebagai sebuah studi ilmiah interdisipliner mengenai bahasa dari sudut pandang komputer, yang menitikberatkan pada rekayasa program aplikasi komputer sehingga komputer mampu melakukan pemrosesan dan pengelolaan terhadap bahasa natural sebagaimana yang biasa dilakukan oleh manusia. Ruang lingkup kajiannya mencakup berbagai bidang kajian yang sama dalam kajian linguistik, mulai dari fonologi, morfologi, leksikografi, sintaksis, semantik dan sebagainya. Hanya saja semua hal tersebut dikaji dari sudut pandang komputer.

Secara epistemologis, keberadaan linguistik komputasi sah disebut sebagai ilmu pengetahuan. Di dunia Barat, ilmu ini telah menjadi isu mutakhir dan telah menjadi bagian penting dalam kajian-kajian akademik di berbagai perguruan tinggi. Namun, di berbagai perguruan tinggi di Indonesia ilmu ini masih menjadi sebuah disiplin ilmu yang langka. Berdasarkan data dari prodi-prodi di berbagai perguruan tinggi di Indonesia, respons terhadap kajian linguistik komputasi cukup bagus dan meningkat. Prodi-prodi di perguruan tinggi umum di bawah Kemendikbud yang menawarkan kajian linguistik komputasi di dalam kurikulumnya naik dari 14,29% menjadi 52,38% dalam waktu 10 tahun terakhir (2011 sampai 2021). Atau naik sebesar 38,10%. Hal ini menunjukkan bahwa kajian linguistik komputasi dianggap sebagai sebuah kajian yang cukup penting di prodi-prodi linguistik. Namun, jika memperhatikan data yang ada, maka masih ada 10 prodi, atau 47,62% prodi yang belum merespons terhadap kajian linguistik sebagai bagian

penting. Sementara, prodi-prodi Bahasa Arab, Bahasa dan Sastra Arab, serta Bahasa Inggris yang berada di bawah Kemenag, masih belum ada satupun yang menjadikan linguistik komputasi sebagai bagian dari kajian linguistik, dan belum ada yang memasukkannya sebagai bagian dari kurikulumnya. Hal ini menjadi tantangan dan pekerjaan rumah (PR) bagi para akademisi linguistik yang perlu direspons, karena kajian linguistik komputasi ini merupakan bagian penting dalam kajian linguistik di era modern ini.

Belum diresponsnya kajian linguistik komputasi di kalangan akademisi linguistik di berbagai perguruan tinggi ini sangat mungkin berkaitan dengan problem basis keilmuan yang mereka miliki di dalam memahami dan mengembangkan disiplin kajian linguistik komputasi. Kompetensi dan keahlian dalam bidang linguistik bisa jadi sudah ada dan dimiliki oleh akademisi di prodi linguistik atau prodi bahasa, sedangkan kompetensi dan keahlian dalam bidang ilmu komputer, umumnya ada pada para akademisi di prodi teknik informatika atau prodi ilmu komputer. Karenanya, akademisi linguistik memiliki problem basis keilmuan berkaitan dengan kompetensi dan keahlian dalam bidang komputer. Untuk merespons tantangan yang muncul dari hadirnya disiplin baru linguistik komputasi ini, perlu upaya dari institusi atau prodi linguistik untuk mendorong para akademisi linguistik mempelajari linguistik komputasi. Hal ini bisa dilakukan melalui pelatihan dalam bidang linguistik. Di samping itu, untuk mentransformasikan dan mengembangkan disiplin linguistik komputasi, akademisi linguistik haruslah didorong untuk membekali diri dengan mempelajari teori komputasi dan berbagai pemrogramannya, lalu diaplikasikan dalam kajian linguistik komputasi. Tanpa kompetensi dan keahlian dalam bidang ilmu komputer, sulit rasanya bagi kalangan akademisi linguistik untuk bisa merespons, memahami, dan mengembangkan disiplin linguistik komputasi.

DAFTAR PUSTAKA

- Audi, Robert. 1999. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. London; New York: Routledge.
- BAN-PT. n.d. "Direktori Hasil Akreditasi Program Studi." Badan Akreditasi Nasional Perguruan Tinggi. Accessed November 26, 2011. https://www.banpt.or.id/?page_id=1903.
- Bolshakov, Igor A., and Gelbukh Alexander. 2004. *Computational Linguistics: Models, Resources, Applications*. Mexico: Ciencia De La Computación.
- Burke, Peter. 1992. *History and Social Theory*. Cambridge: Polity Press.
- Bussmann, Hadumod. 1996. *Routledge Dictionary of Language and Linguistics*. New York: Routledge.
- Carnabuci, Gianluca. 2005. *A Theory of Knowledge Growth*. Amsterdam: Vossiuspers USA.
- Grishman, Ralph. 1994. *Computational Linguistics: An Introduction*. First Edition. Cambridge; Melbourne: Cambridge University Press.
- Hausser, Roland. 2001. *Foundations of Computational Linguistics: Human-Computer Communication in Natural Language*. Berlin; New York: Springer.
- Lacey, Alan. 1996. *Dictionary of Philosophy*. London; New York: Routledge.
- Microsoft. 2002. *The Microsoft Computer Dictionary*. New York: Microsoft Corporation.
- Mitkov, Ruslan. 2004. *The Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.
- Moser, Paul K. 2002. *The Oxford Handbook of Epistemology*. Chicago: Ilionis.
- Musthofa. 2010. "Linguistik Komputer (Computational Linguistic): Model Baru Kajian Linguistik dalam Perspektif Komputer." *Adabiyat Jurnl Bahasa dan Sastra* 9 (December).
- "PDDikti - Pangkalan Data Pendidikan Tinggi." n.d. Accessed June 5, 2021. <https://pddikti.kemdikbud.go.id/prodi>.
- Richards, Jack C., and Richard Schmidt. 2002. *Longman Dictionary of Language Teaching and Applied Linguistics*. Edinburgh: Pearson.
- Ritzer, George. 1996. *Modern Sociological Theory*. New York: The McGraw-Hill Companies.
- Routledge. 2000. *Concise Routledge Encyclopedia of Philosophy*. London, New York: Routledge.
- Salomaa, Arto. 1985. *Computation and Automata*. Cambridge: Cambridge University Press.

- Schubert, Lenhart. 2016. "Computational Linguistics (Stanford Encyclopedia of Philosophy)." In *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/spr2020/entries/computational-linguistics/>.
- Setditjen Dikti. 2020. *Statistik Perguruan Tinggi Tahun 2020 Kemendikbud*. Jakarta: Kemendikbud.
- Toynbee, Arnold J. 1972. *A Study of History*. New York; London: Oxford University Press; Thames and Hudson.
- Verhaar, J.W.M. 2006. *Asas-Asas Linguistik Umum*. 6th ed. Yogyakarta: Gajah Mada University Press.
- Walter, Elizabeth. 2008. "Cambridge Advanced Learner's Dictionary." In *Wikipedia*, 3rd ed. Cambridge: Cambridge University Press.
- Webster, Merriam. 1994. "Merriam-Webster's Collegiate Dictionary." In *Britannica Encyclopaedia*.