

KEBERMAKNAAN SOAL UJIAN NASIONAL BAHASA INDONESIA SMA/MA 2012

Burhan Nurgiyantoro

Fakultas Bahasa dan Seni Universitas Negeri Yogyakarta
e-mail: burhan@uny.ac.id

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi kadar kebermaknaan komponen materi serta ketepatan komponen konstruksi dan bahasa soal Ujian Nasional Bahasa Indonesia (UN BI). Sumber data penelitian adalah soal UN BI SMA/MA 2012 dengan sampel kelompok IPA satu dan IPS tiga set soal. Pengumpulan data dilakukan dengan teknik pembacaan dan pencatatan, sedang analisis data dengan teknik deskriptif kualitatif dengan dibantu statistik deskriptif yang berupa penghitungan frekuensi pemunculan dan persentase. Hasil penelitian sebagai berikut. Pertama, secara umum kadar kebermaknaan komponen materi sudah baik walaupun belum baik sekali. Sebagian besar butir soal sudah melibatkan aspek makna dan bahasa sekaligus sehingga mencerminkan kebutuhan orang berbahasa secara nyata. Kedua, secara umum ketepatan komponen konstruksi dan bahasa juga sudah cukup baik. Sebagian besar butir soal sudah memenuhi tuntutan pembuatan soal yang benar. Namun, secara relatif masih banyak ditemukan kekurangtepatan penggunaan berbagai aspek kebahasaan, bahkan ada yang salah yang menunjukkan kekurangtelitian tim penyusun dan pengulas soal.

Kata kunci: kebermaknaan, ujian nasional, bahasa Indonesia SMA/MA

THE MEANINGFULNESS OF THE INDONESIAN LANGUAGE TEST FOR SMA/MA IN THE NATIONAL EXAMINATION 2012

Abstract

This study aims to identify the degree of the meaningfulness of the materials and the accuracy of the construction and language of the Indonesian Language Test (ILT) in the National Examination (NE). The data sources were the ILT for SMA/MA in the NE 2012 with a sample consisting of one set for the science program and three sets for the social studies program. The data were collected through reading and noting and were analyzed using the quantitative descriptive technique by means of descriptive statistics involving the frequency and percentage calculation. The results are as follows. First, in general the materials are meaningful although they are not maximal yet. Most of the test items take into account of both meaning and language aspects, reflecting the needs for language in actual use. Second, in general the construction and language are moderately accurate. Most of the test items satisfy the criteria for a good test construction. However, there are still inaccuracies in the use of several language aspects, and even there are mistakes, showing test developers' and reviewers' carelessness.

Keywords: meaningfulness, National Examination, Indonesian Language for SMA/MA

PENDAHULUAN

Tagihan terhadap hasil pembelajaran yang diselenggarakan selama jangka

waktu tertentu harus dilakukan. Hal itu dimaksudkan tidak saja untuk mengetahui capaian kompetensi yang berhasil

dicapai oleh peserta didik, melainkan juga keberhasilan program pembelajaran yang diselenggarakan oleh sekolah dan program-program lain yang terkait sebagai satu kesatuan. Walau tampak disederhanakan, hasil capaian peserta didik meraih kompetensi pembelajaran, dalam banyak hal dilihat sebagai hasil keseluruhan program pembelajaran yang dilakukan di sekolah.

Capaian tersebut dapat berupa skor ulangan umum, ulangan umum bersama (UUB), ujian sekolah, dan lain-lain sampai Ujian Nasional (UN). Selama ini pelaksanaan UN masih menjadi isu yang kontroversial. Namun, bagaimanapun UN diperlukan. Ia tidak saja untuk mengukur capaian pembelajaran peserta didik selama belajar dalam satu satuan pendidikan, melainkan juga untuk mengukur efektivitas program pembelajaran yang dilakukan tiap sekolah di berbagai pelosok tanah air dan untuk keperluan perbandingan capaian antarsekolah di berbagai pelosok tersebut.

Sejak era Kurikulum Tingkat Satuan Pendidikan (KTSP) yang sejalan dengan prinsip otonomi daerah, tiap satuan pendidikan di berbagai pelosok diberi kebebasan berkegiatan untuk memajukan dan meraih prestasi. Tujuan KTSP antara lain (i) meningkatkan mutu pendidikan melalui kemandirian dan inisiatif sekolah dalam mengembangkan kurikulum, mengelola dan memberdayakan sumber daya yang ada, dan (ii) meningkatkan kompetensi sehat antarsatuan pendidikan tentang kualitas pendidikan yang dicapai (Mulyasa, 2007:22). Kompetensi sehat dapat dalam berbagai indikator, tetapi salah satunya adalah raih kompetensi yang dibelajarkan, antara lain nilai UN, oleh peserta didik.

Hal itu berlaku untuk berbagai mata pelajaran yang di-UN-kan, termasuk di dalamnya mata pelajaran Bahasa Indonesia. Nilai UN yang kini sudah merupakan gabungan antara nilai UN (60%) dan nilai sekolah (40%) dipandang mewakili capai-

an kompetensi yang dibelajarkan selama peserta didik mengikuti kegiatan pembelajaran. Untuk jenjang SMP dan SMA misalnya, ditempuh selama tiga tahun. Persoalannya adalah apakah soal-soal yang diujikan dapat mewakili kompetensi yang berhasil diraih oleh peserta didik. Apakah peserta didik yang memperoleh nilai tinggi betul-betul menguasai lebih banyak kompetensi daripada peserta didik yang lebih rendah nilainya?

Jika hal itu direduksi khusus untuk mata pelajaran Bahasa Indonesia, pertanyaannya adalah apakah tinggi rendahnya nilai UN Bahasa Indonesia sekaligus mencerminkan kompetensi berbahasa Indonesia peserta didik? Jawaban yang semestinya harus begitu. Namun, pada kenyataannya kondisi itu belum terwujud. Hal itu terlihat misalnya, kompetensi berbahasa Indonesia lulusan SMA/MA baik secara aktif reseptif maupun aktif produktif pada umumnya ternyata tidak sebaik nilai UN-nya. Maka, persoalan yang kemudian muncul adalah mengapa hal itu dapat terjadi?

Selama ini pembelajaran bahasa Indonesia masih terlihat lebih ditekankan pada pembelajaran pada sistem bahasa, bukan pada penggunaan bahasa sebagaimana fungsinya sebagai sarana berkomunikasi. Demikian juga penilaian hasil pembelajaran bahasa Indonesia karena yang demikian lebih mudah. Ujian bahasa Indonesia belum dimaknai sebagai ujian berbahasa Indonesia. Padahal, penguasaan sistem bahasa belum tentu berarti kemampuan mempergunakan bahasa dalam konteks berbahasa. Kompetensi terhadap bahasa belum tentu berarti kompetensi mempergunakan bahasa itu untuk keperluan berkomunikasi secara benar secara kontekstual.

Sejalan dengan pelaksanaan KTSP di dunia pendidikan di Indonesia, yang menekankan capaian kinerja sesuai dengan karakteristik masing-masing mata pelajaran, pelaksanaan pembelajaran di kelas (disarankan) dilakukan dengan

model pembelajaran kontekstual (CTL, *Contextual Teaching and Learning*) (Depdiknas, 2006). Penerapan pembelajaran kontekstual dalam mata pelajaran Bahasa Indonesia menuntut penggunaan bahasa Indonesia secara aktif untuk berkomunikasi. Jadi, penekanan kompetensi yang dibelajarkan adalah kompetensi mempergunakan bahasa, bukan sekadar pengetahuan tentang sistem bahasa. Pengetahuan tentang sistem bahasa penting, tetapi hal itu hanya sebagai sarana untuk menjaga penggunaan bahasa benar. Pengetahuan sistem bahasa penting, tetapi kompetensi berbahasa lebih penting lagi.

Kompetensi mempergunakan bahasa haruslah diukur lewat tagihan yang sesuai, yaitu praktik mempergunakan bahasa itu untuk berkomunikasi dalam konteks tertentu. Bentuk tagihan itu dikenal dengan sebutan tes otentik (*authentic test, authentic assessment*). Pelaksanaan pengukuran dengan tes otentik merupakan salah satu konsekuensi dari pelaksanaan KTSP dengan model pembelajaran kontekstual (Depdiknas, 2006). Dengan kata lain, sejalan dengan pelaksanaan KTSP di dunia pendidikan di Indonesia, pembelajaran yang diselenggarakan di sekolah harus dilakukan dengan model pembelajaran kontekstual dan pengukuran capaian belajar, antara lain, dengan tes otentik. Artinya, penggunaan tes otentik sebagai salah satu alat untuk menagih capaian pembelajaran peserta didik merupakan sebuah keniscayaan.

Tes otentik mengukur capaian kinerja dan sekaligus harus bermakna, menekankan kemampuan peserta didik untuk mendemonstrasikan pengetahuan yang dimiliki secara nyata dan bermakna. Kegiatan penilaian tidak sekadar menanyakan atau menyadap pengetahuan yang telah diketahui pembelajar, melainkan berkinerja secara nyata dari pengetahuan dan keterampilan yang telah dikuasai. Sebagaimana dinyatakan Mueller (2008) tes otentik merupakan suatu bentuk tugas yang menghendaki pembelajar

untuk menunjukkan kinerja di dunia nyata secara bermakna yang merupakan penerapan esensi pengetahuan dan keterampilan.

Pengertian bermakna dalam tes otentik adalah bahwa kinerja yang dilakukan mesti mencerminkan kehidupan nyata secara kontekstual, kinerja yang dibutuhkan dalam kehidupan konkret seperti kebutuhan kompetensi di dunia kerja. Dalam konteks pembelajaran bahasa Indonesia, hal itu menegaskan bahwa pelaksanaan pembelajaran haruslah mengaitkan berbagai bentuk penggunaan bahasa untuk berbagai keperluan di masyarakat. Semakin banyak keterkaitan yang dapat ditemukan peserta didik dalam konteks yang luas, hal itu akan semakin bermakna bagi mereka. Penemuan makna dalam kegiatan pembelajaran adalah ciri utama dari pembelajaran kontekstual (Johnson, 2010:35). Makna itu sendiri dapat diartikan sebagai arti penting dari sesuatu yang dimaksud. Tujuan utama seseorang melakukan sesuatu adalah melihat dan atau menemukan makna yang ada di dalamnya.

Penemuan makna itulah yang seharusnya dilakukan dalam konteks pembelajaran dan tagihan kompetensi berbahasa Indonesia peserta didik. Jika kegiatan pembelajaran dilakukan dengan model pembelajaran kontekstual yang menekankan pencari dan penemuan makna dari suatu kegiatan berbahasa, tagihan yang dimaksudkan untuk mengukur ketercapaian kompetensi juga harus lewat alat yang sesuai. Dalam tagihan kompetensi berbahasa, alat itu berupa tes kinerja, tes perfomansi bahasa yang sekaligus menunjukkan penguasaan terhadap sistem bahasa yang benar penggunaannya sesuai dengan konteks.

Namun, persoalannya adalah tes kinerja berbahasa tidak praktis untuk ujian-ujian yang secara ketat dibatasi oleh waktu. Misalnya, UN yang dibatasi hanya dalam waktu 120 menit, tetapi dimaksudkan untuk mengukur sekian banyak

kompetensi dasar yang dibelajarkan selama waktu tertentu (baca: tiga tahun untuk SMA/MA). Dalam ujian-ujian yang demikian, bagaimanapun tes tradisional bentuk pilihan ganda lebih tepat karena lebih praktis dalam segala hal. Tes otentik yang memrasyaratkan tes kinerja, demonstrasi, atau praktik melakukan sesuatu lebih tepat dilaksanakan dalam penilaian proses. Penilaian proses dilakukan bersamaan dengan proses pembelajaran dan bahkan dapat dijadikan sebagai salah satu strategi pembelajaran.

Keadaan itu sejalan dengan yang dikemukakan Callison (2009), yaitu bahwa tes otentik merupakan sebuah penilaian proses yang di dalamnya melibatkan berbagai kinerja yang mencerminkan bagaimana peserta didik belajar, capaian hasil, motivasi, dan sikap yang terkait dengan aktivitas pembelajaran. Hal juga terkait dengan asumsi bahwa dalam penilaian otentik, ada sekian banyak unjuk kerja yang dapat ditampilkan peserta didik selama berlangsungnya kegiatan pembelajaran yang kesemuanya itu lebih luas dari sekadar ujian tertulis dengan jawaban singkat sebagaimana dalam tes tradisional. Capaian hasil yang diperoleh dari tes otentik langsung dapat ditindaklanjuti dalam pembelajaran selanjutnya.

Persoalan yang kini muncul adalah bagaimana menggabungkan semangat tes otentik yang menekankan kebermaknaan suatu tindak berbahasa dengan tes tradisional bentuk pilihan ganda yang hanya merespon jawaban, tetapi dapat mencakup sekian banyak soal dalam waktu yang relatif singkat. Dengan karakteristik yang memrasyaratkan kinerja, tes otentik tidak dapat dilaksanakan dalam UN (Suwandi, 2012). Namun, UN juga tidak dibenarkan jika hanya mengukur pengetahuan kebahasaan, tanpa nuansa kompetensi berbahasa yang melibatkan makna. Dengan kata lain, soal-soal ujian bahasa Indonesia semacam UN, harus memiliki syarat kebermaknaan. Jadi, apa yang dibelajarkan di sekolah lewat KTSP

dan model pembelajaran kontekstual terwakili dalam UN.

Pertanyaan selanjutnya adalah apakah soal-soal UN Bahasa Indonesia (selanjutnya: UN BI) sudah memenuhi tuntutan kebermaknaan kegiatan berbahasa yang dimaksud? Atau tepatnya, apakah soal-soal UN sudah tidak lagi sekadar menanyakan pengetahuan tentang sistem bahasa dan menekankan kompetensi berbahasa? Selain itu, ketepatan konstruksi dan penggunaan bahasa berpengaruh, maka keduanya perlu juga diteliti. Penelitian ini bertujuan untuk menjawab persoalan-persoalan tersebut. Hal itu dilandasi pertimbangan bahwa bentuk, kandungan, dan misi yang ditanyakan dalam soal-soal UN memengaruhi apa yang dibelajarkan di sekolah. Para guru di sekolah pasti membelajarkan berbagai kompetensi yang akan ditanyakan di UN karena mereka tidak mau dituduh gagal.

METODE

Subjek penelitian ini adalah soal Ujian Nasional Tahun Pelajaran 2011/2012 SMA/MA, Mata Pelajaran Bahasa Indonesia untuk Program Studi IPS dan IPA. Soal UN BI yang berhasil diperoleh adalah tiga set soal untuk program studi IPS (dengan kode: B24, D48, dan E51) dan satu set untuk program studi IPA (A63). Tiap set soal terdiri atas lima puluh butir soal. Keempat set soal tersebut dengan jumlah total dua ratus butir soal dapat dipandang sebagai sampel penelitian karena masih ada set-soal yang lain.

Pengumpulan data dilakukan dengan teknik pembacaan dan pencatatan dengan cermat tiap butir soal dengan mempergunakan instrumen peneliti sendiri. Berbagai kriteria yang dipakai untuk pengumpulan data, misalnya kriteria kebermaknaan suatu soal, telah menjadi bagian pikiran dan intuisi peneliti. Telaah ketepatan butir soal lazimnya mempergunakan tiga kriteria, yaitu materi, konstruksi, dan bahasa (Mardapi, 2008). Komponen materi inilah yang menjadi fokus kebermaknaan pe-

nelitian ini, sedang aspek konstruksi dan bahasa menjadi pelengkap karena juga penting untuk memenuhi tuntutan soal yang baik. Kebermaknaan komponen ini dikategorikan *tinggi* jika menitikberatkan keterlibatan aspek kandungan makna dan bahasa sekaligus, *sedang* jika lebih menitikberatkan bahasa, keduanya mesti secara kontekstual, dan *rendah* jika hanya mencakup sistem bahasa.

Analisis data yang utama mempergunakan teknik deskriptif kualitatif dengan dibantu teknik statistik deskriptif yang berupa penghitungan frekuensi dan persentase pemunculan suatu kategori. Kebermaknaan materi ujian dikelompokkan ke dalam tiga kategori, yaitu tinggi, sedang, dan kurang, sedang komponen konstruksi dan bahasa ke dalam kategori baik, sedang, dan kurang. Analisis deskriptif kualitatif dilakukan lewat langkah-langkah kategorisasi, penyajian data lewat rangkuman hasil penghitungan angka-angka dalam bentuk penghitungan frekuensi dan persentase, dan pembuatan inferensi yang dimaksudkan untuk memaknai dan menyimpulkan temuan-temuan.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Hasil penelitian yang mencakup dua kelompok disajikan ke dalam dua tabel rangkuman, yaitu yang terkait dengan kadar kebermaknaan soal-soal UN BI untuk komponen materi dan kadar ketepatan untuk komponen konstruksi dan bahasa. Pemunculan angka-angka tabel sebagai konsekuensi kategorisasi data terhadap keempat perangkat soal ujian yang diteliti.

Pembedaan soal UN berdasarkan materi kebahasaan dan kesastraan juga ditunjukkan karena hal itu sering dipertanyakan pembobotannya masing-masing. Misalnya, berapa banyak atau persen masing-masing butir-butir soal bahasa dan sastra. Hal itu dapat dimengerti karena muara akhir kedua aspek pembelajaran tersebut tidak sama.

Pembahasan

Butir Soal UN Bahasa versus Sastra

Dalam kurikulum sekolah menengah (KTSP 2006/2007: SMP, SMA/MA, SMK) yang dipergunakan di dunia pendidikan di Indonesia dewasa ini mata pelajaran yang semula bernama Bahasa dan Sastra Indonesia dipersingkat menjadi Ba-

Tabel 1. Kadar Kebermaknaan Komponen Materi Butir Soal UN Bahasa Indonesia SMA/MA 2012

No.	Perangkat Soal	Kadar Kebermaknaan			Keterangan
		Tinggi	Sedang	Rendah	
1.	A63 IPA				
	Bahasa :	22 (44%)	15 (30%)	1 (2%)	Bahasa: 37 (74%)
	Sastra :	10 (20%)	2 (4%)	1 (2%)	Sastra : 13 (26%)
2.	B24 IPS				
	Bahasa:	16 (32%)	14 (28%)	1 (2%)	Bahasa: 31 (62%)
	Sastra :	11 (22%)	8 (16%)	-	Sastra : 19 (38%)
3.	D48 IPS				
	Bahasa:	15 (30%)	16 (32%)	1 (2%)	Bahasa: 32 (64%)
	Sastra :	10 (20%)	8 (16%)	-	Sastra : 18 (36%)
4.	E51 IPS				
	Bahasa:	15 (30%)	18 (36%)	1 (2%)	Bahasa: 34 (68%)
	Sastra :	10 (20%)	6 (12%)	-	Sastra : 16 (32%)

hasa Indonesia. Penghilangan “Sastra” sebenarnya tidak menjadi masalah asal tidak dimaknai sebagai tidak pentingnya pembelajaran kesastraan. Bagaimanapun, dalam wacana peningkatan muatan pendidikan karakter di Indonesia dewasa ini, konten pembelajaran sastra amat dibutuhkan. Hal itu disebabkan muara akhir pembelajaran sastra adalah (pembentukan) afektif, sedang pembelajaran bahasa lebih kognitif. Upaya pembentukan karakter itu bahkan juga secara strategis dapat juga dilakukan lewat ujian-ujian berbahasa yang berkarakter otentik (Abidin, 2012). Hal itu semua menunjukkan pentingnya, strategi, prosedur, dan konten evaluasi untuk mencapai kompetensi dan atau tujuan tertentu.

Namun, tidak menutup kemungkinan bahwa penghilangan tersebut juga dimaknai sebagai tidak (dianggap) pentingnya pembelajaran sastra. Atau, mini-

mal tidak sepeenting pada kurikulum sebelumnya. Jika demikian halnya, itu patut disayangkan. Jika penghilangan itu dengan alasan pembelajaran sastra dilakukan lewat saluran empat kompetensi berbahasa, hal itu lebih dapat diterima. Jadi, dalam hal ini sastra sekadar dipandang sebagai konten pembelajaran saja. Faktanya dalam UN BI konten sastra juga dimunculkan dengan persentase yang cukup baik.

Dalam UN BI SMA/MA 2012 bobot soal kesastraan untuk kelompok IPS berada di kisaran persen 30-an (38%, 36%, dan 32%), sedang IPA persen 20-an (26%). Persoalan baru yang muncul kini adalah mengapa jumlahnya tidak sama dan bukan ada atau tidak butir-butir tentang kesastraan. Hal terjadi pada soal UN kesastraan kelompok IPS yang jumlah soalnya berbeda. Padahal, kisi-kisi pengujian yang dipergunakan untuk membuat soal

Tabel 2. Kadar Ketepatan Komponen Konstruksi dan Bahasa Butir Soal UN Bahasa Indonesia SMA/MA 2012

No.	Perangkat Soal	Kadar Ketepatan				Keterangan
		Baik	Sedang	Rendah		
1.	A63 IPA					
	Konstruksi:	45 (90%)	5 (10%)	-	Konstruksi	Bahasa
	Bahasa :	41 (82%)	9 (18%)	-	Baik: 45 (90%)	Baik: 41 (82%)
					Sedang+Rendah: 5 (10%)	Sedang+rendah: 9 (18%)
2.	B24 IPS					
	Konstruksi:	48 (96%)	2 (4%)	-	Konstruksi	Bahasa
	Bahasa :	43 (86%)	7 (14%)	-	Baik: 48 (96%)	Baik: 43 (86%)
					Sedang+Rendah: 2 (4%)	Sedang+rendah: 7 (14%)
3.	D48 IPS					
	Konstruksi:	48 (96%)	1 (4%)	1 (4%)	Konstruksi	Bahasa
	Bahasa :	38 (76%)	12 (24%)	-	Baik: 48 (96%)	Baik: 38 (76%)
					Sedang+Rendah: 2 (4%)	Sedang+rendah: 12 (24%)
4.	E51 IPS					
	Konstruksi:	44 (88%)	5 (10%)	1 (2%)	Konstruksi	Bahasa
	Bahasa :	44 (88%)	6 (12%)	-	Baik: 44 (88%)	Baik: 44 (88%)
					Sedang+Rendah: 6 (12%)	Sedang+rendah: 6 (12%)

pasti sama. Keadaan semacam ini dapat diperkirakan karena tim penyusun soal UN BI tidak setia pada kisi-kisi yang dijadikan acuan. Reviuer soal mungkin juga kurang teliti. Walau tidak fatal, hal tersebut tetap saja patut disayangkan karena soal menjadi tidak lagi setara di antara set-set yang disediakan.

Kebermaknaan Soal UN BI

Sebenarnya, tuntutan kebermaknaan tidak hanya permasalahan penilaian, melainkan juga permasalahan pembelajaran. Penilaian hanya mengukur kompetensi yang dibelajarkan. Hal ini menjadi persoalan pembelajaran bahasa secara umum dan universal. Penekanan pembelajaran bahasa pada kebermaknaan itulah yang memacu munculnya berbagai pendekatan pembelajaran, seperti pendekatan komunikatif, pragmatik, kontekstual, dan lain-lain. Demikian juga tagihan terhadap hasil pembelajaran bahasa. Penilaian (baca: UN BI) haruslah bermakna, atau paling tidak menekankan kebermaknaan. Hal itulah yang kemudian memacu munculnya berbagai teknik pengukuran kompetensi berbahasa yang menekankan kinerja (misalnya: McNamara, 1997; Brown, 2004, 2007; Mueller, 2008; Calison, 2007, 2009; Wahyuni, 2009). Intinya, penilaian kinerja dan bermakna itulah yang kini menjadi kecenderungan sesuatu yang harus dilakukan.

UN BI juga tidak lepas dari tuntutan tersebut. Artinya, butir-butir soal UN BI harus juga memenuhi kualifikasi kebermaknaan walau sekadar menuntut kompetensi berbahasa aktif reseptif saja, yaitu lewat membaca dan memilih jawaban. Tuntutan kinerja aktif produktif berbahasa yang menuntut kompetensi mengonstruksi sendiri jawaban, misalnya dalam bentuk berbicara dan menulis, diasumsikan sudah dilaksanakan pada penilaian proses sebelum UN diselenggarakan. Keadaan itu terlihat dari "hak" dan atau pertimbangan guru untuk ikut meluluskan peserta didik sebesar 40%.

Hasil penelitian terhadap soal-soal UN BI SMA/MA 2012, baik untuk kelompok IPA maupun IPS menunjukkan bahwa sebagian besar butir soal memenuhi tuntutan kebermaknaan. Dalam penelitian ini kebermaknaan dikategorikan ke dalam tiga tingkatan, yaitu tinggi, sedang, dan rendah. Baik soal yang berkategori tinggi maupun sedang memrasyaratkan konteks. Kejelasan konteks akan secara pasti membimbing menunjukkan bentuk kebahasaan yang tepat pula. Hasil penelitian menunjukkan bahwa butir-butir soal UN BI yang berangkat dari konteks cukup mendominasi keseluruhan soal, sedang butir-butir soal yang tidak mendasarkan diri pada konteks hanya beberapa persen saja (lihat Tabel 1). Dilihat dari sudut ini, dapat dikatakan bahwa soal-soal UN BI sudah bagus walau masih dapat dibuat lebih bagus lagi.

Tingkat kebermaknaan soal dinyatakan tinggi jika untuk dapat menjawab dengan benar butir soal yang bersangkutan diperlukan referensi bahasa dan makna sekaligus. Hal itu sejalan dengan pandangan tentang bahasa bahwa yang utama dan pertama berfungsi sebagai sarana komunikasi (Brown, 2004, 2007). Dalam konteks berkomunikasi yang sebenarnya, konten pembicaraan dan bahasa sama-sama dibutuhkan. Bahkan, tampaknya aspek konten (baca: isi pembicaraan) sering dianggap lebih penting. Maka, tagihan kompetensi berbahasa harus menyangkut kedua aspek itu. Dengan kata lain, soal ujian bahasa yang mencakup kedua aspek dalam konteks yang jelas dimaknai sebagai memunyai kadar kebermaknaan yang tinggi.

Butir-butir soal yang berkadar kebermaknaan tinggi misalnya adalah soal-soal yang terkait dengan tagihan terhadap kandungan isi teks-teks bacaan, baik teks sastra maupun nonsastra, membaca, menafsirkan, dan memaknai tabel, surat, atau register bahasa yang lain (seperti memahami, mengidentifikasi, menafsirkan, menyimpulkan). Kegiatan berbahasa yang

demikian juga ditemukan dan dibutuhkan dalam kehidupan nyata, maka kegiatan itu sudah bernuansakan penilaian otentik. Mengapa hanya “bernuansakan”? Hal itu disebabkan penilaian otentik yang benar-benar otentik lebih berwujud kegiatan aktif produktif daripada “sekadar” aktif reseptif. Namun, untuk ujian-ujian yang harus menagih banyak aspek (bahan) dengan waktu yang terbatas, hal itu sudah cukup baik.

Contoh butir-butir yang dimaksud adalah sebagai berikut.

(1) Disediakan satu alenia sebuah wacana yang tidak terlalu panjang kemudian diikuti dua buah soal yang menanyakan makna wacana (misal: D48 IPS, No.19 dan 20):

19. Opini penulis dalam tajuk tersebut adalah ...

- A. Bank dunia harus turun tangan mengatasi masalah BBM bersubsidi di tanah air ini.
- B. Penghematan anggaran subsidi harus dialokasikan untuk anggaran belanja pembangunan infrastruktur.
- C. Pemerintah hendaknya menyadari ancaman pembengkakan subsidi di tanah air ini.
- D. Pemerintah sebaiknya menaikkan harga BBM bersubsidi disertai kompensasi bagi masyarakat miskin.
- E. Penghematan anggaran subsidi BBM digunakan untuk peningkatan belanja negara.

20. Opini permasalahan dalam tajuk tersebut ditujukan kepada

- A. Bank Dunia
- B. Perguruan tinggi
- C. pemerintah
- D. pengusaha
- E. masyarakat

(2) Disediakan sebuah tabel konkret dan mutakhir kemudian diikuti dua pertanyaan yang terkait dengan isi (misal: A63 IPA, no.6 dan 7):

6. Pernyataan yang sesuai dengan tabel tersebut adalah ...

- A. Kenaikan harga tertinggi terjadi pada tahun 2006-2007.
- B. Kenaikan harga tertinggi terjadi pada tahun 2007-2008.
- C. Kenaikan harga tertinggi terjadi pada tahun 2008-2009.
- D. Setiap tahun harga elpiji mengalami kenaikan.
- E. Kenaikan harga elpiji tahun 2007-2008 sangat drastis.

7. Simpulan yang tepat berdasarkan tabel tersebut adalah ...

- A. Harga elpiji setiap tahun mengalami kenaikan.
- B. Kenaikan sangat drastis terjadi pada tahun 2008-2009.
- C. Kenaikan harga elpiji terjadi setiap tahun, kecuali pada tahun 2005-2006.
- D. Kenaikan harga elpiji setiap tahun sangat variatif.
- E. Kenaikan harga elpiji setiap tahun dipengaruhi oleh produksi elpiji.

Pemilihan ketepatan butir-butir soal yang pertimbangannya membutuhkan input konten dan bahasa adalah soal yang terbaik karena mencerminkan fungsi dan kebutuhan orang berbahasa secara nyata. Jenis butir soal yang demikianlah yang seharusnya menjadi karakteristik ujian kompetensi berbahasa (Indonesia) sehingga peserta uji tidak hanya berpikir tentang bahasa secara eksklusif. Namun, pernyataan ini tidak perlu dimaknai sebagai pengabaian pentingnya sistem bahasa. Ketepatan penggunaan bahasa secara benar secara sistem penting, namun kondisi itu harus selalu dikaitkan dengan kebutuhan orang berbahasa, bukan secara terpisah (lihat juga Wahyuni, 2009).

Kondisi itu juga berlaku untuk butir-butir soal yang dimaksudkan mengukur kompetensi bersastra. Butir-butir soal yang berkadar kebermaknaan tinggi adalah yang mengukur kompetensi bersastra, bukan sistem tentang sastra. Jadi, soal

mesti berangkat dari sebuah teks kesastraan secara konkret dengan pertanyaan-pertanyaan yang terkait dengan kandungan makna di dalamnya.

Namun, tidak sedikit juga soal-soal UN BI yang lebih ditekankan mengukur ketepatan penggunaan komponen bahasa Indonesia dalam sebuah konteks. Misalnya, ketepatan kalimat, diksi, ungkapan, ejaan, dan lain-lain. Pertanyaan ujian pada umumnya terkait dengan kompetensi membetulkan atau mengganti aspek kebahasaan yang salah, mengenali, dan menempatkan bentuk tertentu, dan lain-lain. Hal itu berarti pertimbangan ketepatan aspek konten tidak penting atau tidak dipentingkan. Bahwa butir-butir soal itu berangkat dari sebuah konteks, maka sudah memenuhi persyaratan kontekstual, fakta itu merupakan poin yang baik. Untuk menjawab dengan benar sebuah butir soal, dengan demikian, masih diperlukan pertimbangan lain selain aspek kebahasaan semata.

Walau level kebermaknaannya tidak setinggi keadaan butir soal seperti di atas, pertanyaan atau tugas yang demikian juga mencerminkan kebutuhan orang berbahasa, yaitu tuntutan bahwa orang harus mempergunakan bahasa secara tepat sesuai dengan konteks. Oleh karena itu, jenis butir-butir soal UN BI yang berciri demikian, dikategorikan memiliki kadar kebermaknaan tingkat *sedang*. Persyaratan untuk dapat menjawab dengan benar suatu butir, "hanya" memrasyaratkan pengetahuan sistem bahasa dan kurang menekankan pentingnya komponen konten kegiatan berbahasa. Butir-butir soal UN BI yang demikian, masih tergolong baik, maka tidak ada salahnya tetap dipertahankan. Hanya saja, jumlahnya perlu dibatasi karena butir soal yang terbaik adalah yang memiliki kadar kebermaknaan tinggi.

Tabel 1 di atas memperlihatkan bahwa butir-butir soal UN BI yang berkategori demikian masih relatif tinggi, bahkan untuk set soal D48 IPS dan E51 IPS lebih

banyak daripada soal yang kadar kebermaknaannya tinggi. Hal itu menunjukkan bahwa tim penyusun UN BI masih mempertimbangkan lebih pentingnya ketepatan aspek bahasa daripada ketepatan aspek bahasa dan konten sekaligus. Namun, sekali lagi ada hal yang mengusik: mengapa untuk set soal B24 IPS lebih banyak soal yang kadar kebermaknannya tinggi, padahal pasti ketiga set soal itu disusun berdasarkan kisi-kisi pengujian yang sama? Sekali lagi, kondisi itu mungkin terjadi karena tim penyusun dan reviuer kurang mematuhi rambu-rambu kisi-kisi pengujian.

Contoh butir-butir yang dimaksud adalah sebagai berikut.

- (1) Disediakan satu alenia sebuah wacana yang di dalamnya terdapat kata-kata tugas yang salah dengan dicetak miring. Peserta uji diminta membetulkan kata-kata salah itu dengan memilih sejumlah pilihan yang disediakan (misal: B24 IPS, No.12).
12. Kata penghubung yang tepat untuk memperbaiki kata penghubung yang bercetak miring pada paragraf tersebut adalah
 - A. atau, tetapi, sebab
 - B. yaitu, bahkan, bahwa
 - C. serta, yakin, dan
 - D. dan, serta, jika
 - E. karena, misalnya, apabila
- (2) Disediakan kalimat-kalimat topik untuk berpidato, kemudian peserta uji diminta mengenali kalimat jenis tertentu (misal: B24, no. 13).
13. Kalimat persuasif yang sesuai dengan topik tersebut terdapat pada nomor
 - A. (1) dan (2)
 - B. (1) dan (3)
 - C. (2) dan (3)
 - D. (3) dan (4)
 - E. (4) dan (5)

Sebenarnya, soal-soal yang berkategori kebermaknaan tinggi dan sedang

juga bergradasi. Artinya, butir-butir soal tidak sama kadar tinggi dan sedangnya walau sama-sama dapat dikategorikan dalam posisi yang sama. Contoh kedua soal di atas (nomor 12 dan 13) misalnya, level kebermaknaannya tinggi nomor 12. Butir tersebut mengandung tuntutan penempatan bentuk tertentu secara tepat konteks, jadi kinerja kognitif yang dibutuhkan lebih tinggi, sedang nomor 13 hanya menuntut pengenalan saja.

Penelitian ini menunjukkan bahwa butir-butir soal UN BI yang kadar kebermaknaannya rendah jumlahnya sedikit (lihat Tabel 1). Keadaan itu dapat dipandang sebagai salah hal yang baik untuk soal UN BI. Sebuah butir soal dikategorikan berkategori kebermaknaan rendah jika soal itu hanya mengukur sistem atau sastra saja. Indikatornya yang terlihat adalah soal-soal yang tidak dibuat berdasarkan konteks. Jadi, untuk dapat menjawab dengan benar soal-soal tersebut hanya memerlukan satu pertimbangan, yaitu pertimbangan ketepatan (sistem) bahasa. Soal-soal yang berciri demikian banyak dijumpai pada ujian-ujian bahasa tempo dulu ketika pembelajaran bahasa masih berangkat dari pandangan strukturalisme.

Butir-butir tersebut mungkin “hanya” mengukur unsur pembentuk kalimat, makna kata, atau ejaan saja yang mungkin berupa tuntutan untuk mengenali, mengidentifikasi, membedakan, menunjukkan, dan lain-lain. Soal bisa saja berangkat dari konteks wacana, tetapi ia hanya dipakai sebagai tempat terdapatnya unsur yang dimaksud. Dalam soal UN BI contoh yang ditemukan umumnya terkait dengan penulisan ejaan. Ketika pembelajaran bahasa menekankan pentingnya pembelajaran kontekstual dengan kompetensi berbahasa sebagai target akhir capaian sebagai tuntutan dewasa ini, butir-butir soal yang hanya mengukur kompetensi (sistem) bahasa dan sastra tidak lagi dapat diterima. Capaian kompetensi itu memang diperlukan, tetapi ia hanya

penting dalam kaitannya dengan kegiatan berbahasa secara bermakna.

Penelitian kebermaknaan ini dilakukan pada komponen materi. Dilihat dari segi materi yang diujikan, tampaknya butir-butir soal yang dibuat tidak semuanya memenuhi tuntutan materi yang sesuai dengan satuan pendidikan SMA/MA. Terdapat sejumlah butir soal yang terlihat “hanya seperti itu”, terlalu sederhana untuk level sekolah menengah atas. Misalnya, butir soal yang “hanya” mengukur kompetensi pengenalan bentuk, penggunaan kata tugas (ada banyak soal untuk ini), merampungkan pantun, memilih penutup surat, dan lain-lain tampak terlalu sederhana karena konteks yang dibuat juga sederhana.

Konstruksi dan Bahasa

Aspek konstruksi dan bahasa dalam sebuah butir soal memengaruhi ketepatan bentuk dan keterbacaan yang selanjutnya memengaruhi kepastian pilihan jawaban benar oleh peserta uji. Keterbacaan soal-soal ujian harus dipandang sebagai sesuai yang penting, maka harus diusahakan keterpenuhannya. Selain aspek materi, kedua aspek ini juga menentukan ketepatan sebuah soal sebagai alat uji yang baik, maka keduanya juga harus baik (Mardapi, 2008). Konsekuensi dalam penelitian ini adalah keduanya juga harus diteliti walau bukan merupakan fokus utama.

Konstruksi terkait dengan bagaimana sebuah butir soal dibuat yang antara lain mencakup pembuatan pokok soal dan opsi jawaban. Pokok soal dan opsi misalnya, harus dirumuskan dengan jelas, tidak membingungkan, dan opsi harus segramatikal dengan pokok soal. Artinya, opsi harus cocok dengan pertanyaan pokok soal dan homogen. Panjang opsi harus kurang lebih sama, antarbutir tidak boleh saling menggantungkan, dan cara penulisan angka harus mengikuti ketentuan tertentu. Aspek bahasa terkait dengan penggunaan bahasa yang baku, komunikatif, dan makna tidak ambigu.

Hal-hal tersebut merupakan persyaratan yang harus terpenuhi. Jika tidak terpenuhi, dikhawatirkan jika peserta didik salah menjawab, hal itu lebih disebabkan perintahnya yang tidak jelas daripada penguasaan kompetensi yang diujikan.

Hasil penelitian menunjukkan bahwa secara umum aspek konstruksi dan bahasa soal UN BI sudah lebih banyak yang berkategori baik. Artinya, butir-butir soal itu sudah memenuhi tuntutan kualifikasi soal ujian yang baik. Jika kedua aspek itu dibandingkan, unsur ketidaktepatan bahasa jauh lebih banyak daripada konstruksi. Hal itu dapat dipandang sebagai sesuatu yang ironis karena soal yang diteliti adalah soal ujian bahasa Indonesia. Soal ujian BI seharusnya mempergunakan BI yang benar, baik dari diksi, struktur, kandungan makna yang ditunjuk, maupun ejaan. Sebagaimana dikemukakan sebelumnya, keadaan itu menunjukkan bahwa tim penyusun soal UN BI kurang teliti.

Mengidentifikasi kekurangtepatan aspek konstruksi mungkin lebih mudah karena lebih kasat mata, maka kekurangtepatan aspek ini juga tidak banyak. Unsur kekurangan itu lebih banyak terlihat pada fakta bahwa panjang opsi yang sering "tidak kurang lebih sama". Tuntutan konstruksi soal yang baik antara lain adalah panjang opsi kurang lebih sama. Hal inilah yang terlihat kurang dapat dipenuhi pada sejumlah butir soal. Kondisi konstruksi butir soal yang lain seperti pokok soal jelas, opsi gramatikal dengan pokok soal dan homogen terlihat sudah terpenuhi.

Keadaan itu terlihat berbeda dengan aspek bahasa yang menunjukkan adanya kesalahan atau ketidaktepatan yang lebih banyak dan lebih beragam. Ketidaktepatan itu antara terlihat pada perintah melakukan sesuatu, penunjukan wacana, makna ambigu atau bahkan salah logika, sampai dengan ejaan. Dalam penelitiannya terhadap soal UN BI SMP 2010/2011, Suwandi (2012) juga menemu-

kan berbagai bentuk kekurangtepatan aspek bahasa. Kesalahan-kesalahan itu misalnya, ditemukan pada kesalahan struktur kalimat, morfologi, ejaan, dan juga konstruksi.

Sebuah butir soal harus tidak "ternodai" oleh kesalahan-kesalahan "kecil" semacam itu karena UN adalah dokumen resmi negara. Oleh karena itu, butir-butir yang memiliki kesalahan-kesalahan tersebut tidak dapat dinyatakan berkategori *baik*, melainkan *sedang* atau bahkan *rendah* tergantung tingkat keparahan kesalahan. Butir yang berkategori baik haruslah terhindar dari kesalahan-kesalahan baik yang menyangkut aspek bahan, konstruksi maupun bahasa.

Contoh data penelitian yang menunjukkan adanya penggunaan bentuk-bentuk bahasa yang tidak tepat itu tidak hanya terdapat pada butir soal, melainkan juga terdapat dalam penunjukan wacana dan perintah. Keadaan itu bahkan ditemukan berkali-kali. Contoh yang dimaksud sebagai berikut.

- (1) Kutipan novel **tersebut** untuk soal nomor 21 dan 22. Bacalah dengan seksama! (D48 IPS)
- (2) Kutipan novel **tersebut** untuk soal nomor 21 dan 22. Bacalah dengan seksama! (A63 IPA)

Penggunaan kata "tersebut" pada perintah itu salah karena kutipan yang dimaksud ditempatkan sesudah perintah itu, bukan sebelumnya. Jadi, yang benar mestinya adalah "berikut" atau "di bawah". Hal itu terjadi mungkin karena penulis soal hanya memfoto kopi perintah pada soal sebelumnya yang menempatkan teks di atas perintah tanpa mencermati penggunaannya. Selain itu, secara umum perintah itu juga tidak jelas. Maksudnya, kutipan itu dimaksudkan untuk "menganalisis" soal yang diberikan. Jadi, perintah itu yang tepat, misalnya berbunyi: "Kutipan novel di bawah ini untuk menjawab butir soal nomor 21 dan 22". Peserta didik memang masih dapat memahami maksudnya, namun kesalahan "kecil"

semacam itu seharusnya tidak perlu terjadi pada ujian yang ser resmi UN.

Belum lagi ternyata soal-soal itu sama walau untuk menguji peserta dari kelompok IPA dan IPS. Penelitian ini juga menemukan sejumlah soal yang sama untuk set-set yang berbeda. Keadaan ini menunjukkan kemalasan tim penyusun membuat soal yang berbeda dan seharusnya juga tidak perlu terjadi. Apa gunanya UN BI dibuat ke dalam sejumlah perangkat jika di dalamnya terdapat butir-butir soal yang persis sama. Sebenarnya, hal itu mudah disiasati dengan mengambil wacana yang berbeda walau yang ditanyakan hal yang tidak berbeda.

Temuan yang agak “memalukan” adalah banyaknya ejaan yang salah. Walaupun kesalahan itu tidak dalam rangka untuk dijawab benar-salahnya, keadaan itu tetap saja salah. Padahal, konteksnya adalah UN BI yang amat resmi. Contohnya, penulisan judul novel yang tidak dicetak miring, melainkan dengan diapit tanda kutip (“—”). Keempat set UN BI yang diteliti pada butir nomor 29 menulis judul salah. Soal set A63 dan B24 menulis “... novel *Memory in Song*”..., sedang soal set D48 dan E51 menulis “Novel “Bekisar Merah...””. Keadaan itu belum lagi mempertimbangkan “dosa” adanya soal yang sama untuk dua set yang berbeda. Yang terlihat janggal, penulisan novel *Memory in Song* tidak konsisten, yang satu dicetak miring, sedang yang lain diapit dua tanda kutip.

Penelitian ini bahkan menemukan dua butir soal yang salah yang dapat dipandang sebagai kesalahan bahasa. Yang pertama terkait dengan soal sastra (D48 nomor 13) dan yang kedua penulisan surat (E52, nomor 42). Yang pertama terkait dengan logika, sedang kedua ketepatan menyalin. Contoh kasus pertama.

15. Cermati pantun berikut!

Pergi berlibur ke pantai
Pulanginya naik odong-odong
Jadilah anak yang pandai

...

Larik yang tepat untuk melengkapi pantun tersebut adalah

- A. Jangan selalu tolong-menolong
- B. Perangai baik dan tidak sombong
- C. Punya teman suka menolong
- D. Jangan bicara omong kosong
- E. Punya otak janganlah kosong

Opsi A soal di atas jelas salah, maka ia tidak akan dipilih oleh peserta didik yang mengerti pantun. Pantun berisi nasihat dan nasihat pasti terkait dengan hal-hal yang baik. Opsi A justru menampilkan isi yang salah dan itu bertentangan dengan semangat (logika) pantun nasihat. Selain itu, karena opsi yang lain bernada positif, maka opsi A tersebut menjadi tidak homogen. Ada kesan penulis soal hanya asal membuat opsi dengan penekanan pada persajakan akhir dan mengabaikan logika dan semangat pantun. Sekali kesalahan-kesalahan “kecil” semacam itu yang dapat berubah menjadi “besar” seharusnya tidak terjadi.

SIMPULAN

Penelitian sampai pada kesimpulan bahwa UN BI Tahun 2012 untuk SMA/MA kelompok IPA dan IPS yang terkait dengan kadar kebermaknaan serta ketepatan aspek konstruksi dan bahasa sebagai berikut. *Pertama*, secara umum kadar kebermaknaan komponen materi sudah baik walau belum baik benar. Sebagian besar butir soal sudah melibatkan aspek makna dan bahasa sekaligus sehingga mencerminkan kebutuhan orang berbahasa secara nyata. Fakta bahwa relatif masih banyak butir soal yang menekankan aspek berbahasa saja walau secara kontekstual, hal itu menunjukkan bahwa bayang-bayang tes bahasa masih kuat yang justru dapat mengurangi kadar kebermaknaannya. Butir soal yang berkadar kebermaknaan rendah masih juga ditemukan walau hanya beberapa butir soal.

Kedua, secara umum ketepatan komponen konstruksi dan bahasa juga sudah cukup baik. Sebagian besar butir soal su-

dah memenuhi tuntutan pembuatan soal yang benar. Namun, secara relatif masih banyak ditemukan kekurangtepatan penggunaan berbagai aspek kebahasaan, bahkan ada yang salah. Selain itu, juga ditemukan sejumlah butir yang sama di antara set-set yang berbeda. Keadaan itu menunjukkan kekurangtelitian dan mungkin juga "kemalasan" tim penyusun dan reviuer soal dalam mengembangkan butir-butir soal, termasuk kepatuhan pada kisi-kisi pengujian.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada sejawat yang membantu penelitian dan penulisan artikel ini hingga bentuknya yang ada di hadapan pembaca, baik yang berupa saran, *sharing*, maupun yang lain-lain. Secara khusus ucapan terima kasih kepada kawan yang memberikan kepada saya soal-soal UN Bahasa Indonesia yang kemudian dijadikan subjek penelitian. Namun demikian, jika terdapat berbagai kekurangan, hal itu semata-mata menjadi tanggung jawab saya sebagai peneliti.

DAFTAR PUSTAKA

- Abidin, Yunus. 2012. "Model Penilaian Otentik dalam Pembelajaran Membaca Pemahaman berorientasi Pendidikan Karakter", dalam *Jurnal Pendidikan Karakter*. Tahun II, No. 2, hlm 164-178.
- Brown, H. Douglas. 2004. *Language Assessment, Principles and Classroom Practices*. San Francisco: Longman.
- Brown, Douglas H. 2007. *Teaching by Principles, an Interactives Approach to Language Paedagogy*. New York: Pearson Seducation.
- Callison, Daniel. 2007. *Authentic Assessment*. Chicago: American Library Association.
- Callison, Daniel. 2009. "Authentic Assessment" dalam *American Assosiation of School Librarians*. <http://www.ala.org/ala/mgrps/divs/aasl/aaslpubsand-journals/slmrb/editorschoiceb/info-power/selctcallison85.cfm>, diunduh 3 Oktober 2009.
- Depdiknas. 2006. *Pendekatan Kontekstual*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Direktorat Pendidikan Menengah Pertama.
- Johnson, Elaine B. 2010. *Contextual Teaching and Learning, Menjadikan Kegiatan Belajar-Mengajar Mengasyikkan dan Bermakna*. tanpa kota: MLC.
- Mardapi, Djemari. 2008. *Teknik Penyusunan Instrumen Tes dan Nontes*. Jogjakarta: Mitra Cendekia.
- McNamara, Tim. 1997. *Measuring Second Language Performance*. London & New York: Longman.
- Mueller, John. 2008. *Authentic Assessment Toolbox*. North Central College, Naperville, <http://jonathan.mueller.faculty.noctrl.edu/toolbox/index.htm>, diunduh 27 Agustus 2008.
- Mulyasa, E. 2007. *Kurikulum Tingkat Satuan Pendidikan*. Bandung: Ramaja Rosdakarya.
- Suwandi, Sarwiji. 2012. "Penilaian Otentik untuk Ujian Nasional Mata pelajaran Bahasa Indonesia (Sebuah Tantangan)", disampaikan dalam *Seminar Nasional Mendamba Ujian Nasional yang Komunikatif dan Otentik*. Universitas Muhammadiyah Surakarta.
- Wahyuni, Sri. 2009. *Pengembangan Model Asesmen Otentik dalam Pembelajaran Keterampilan Berbahasa Indonesia Lisan di SMA*. Malang: Disertasi Program Pascasarjana Universitas Negeri Malang.