

**MULTIMODALITY: MENAFSIR VERBAL, MEMBACA GAMBAR,  
DAN MEMAHAMI TEKS**

**Budi Hermawan**

Jurusan Pendidikan Bahasa dan Sastra Inggris, FPBS UPI

Korespondensi: Jl. Dr. Setiabudhi 229 Bandung 40154

Pos-el: [budi\\_bee@yahoo.com](mailto:budi_bee@yahoo.com)

**Abstrak**

*Tulisan ini ditujukan khususnya untuk memenuhi dua tujuan utama. Pertama, tulisan ini memperkenalkan dan menjelaskan multimodality sebagai sebuah 'prosedur analisa' yang harus digunakan untuk menganalisa teks yang menggunakan lebih dari satu semiotic mode, khususnya yang menggunakan mode verbal dan mode gambar atau imej secara bersamaan dalam sebuah kesempatan penyampaian makna. Kedua, tulisan ini menjelaskan langkah-langkah teknis prosedur analisa multimodality yang dapat digunakan untuk menganalisa teks seperti tersebut dan memberikan contoh penggunaan langkah analisa. Dengan demikian, tulisan ini juga mengeksplorasi manfaat yang dapat diperoleh dari penggunaan 'prosedur analisa' ini untuk menganalisa teks. Tulisan ini mendukung argumen yang ditawarkan diantaranya oleh Kress dan van Leeuwen (2006), dan Machin dan Myer (2012), yang menyakini bahwa pesan yang disampaikan dengan semiotic mode berbeda secara bersamaan (verbal dan imej) dalam sebuah teks tidak dapat dianalisa hanya dengan alat analisa linguistik saja, tetapi mengharuskan dua alat analisa yang berbeda yaitu linguistics, dan image analysis tool seperti reading image yang saling mendukung menuju pemahaman makna yang lebih menyeluruh. Bentuk hubungan yang tercipta dari verbal dan imej sebagai dua semiotic mode berbeda dalam sebuah teks juga dipaparkan dalam tulisan ini.*

**Kata-kata kunci:** multimodality, semiotic mode, teks, mode verbal, mode imej

**Abstract**

*This article was especially to serve two main purposes; firstly, it introduces and elaborates multimodality as a 'procedure of analysis' to use to analyze texts which use more than one semiotic mode, particularly those that use verbal and images at the same time. Secondly, this article both explicates technical steps of the procedure that can be used to analyze given texts, and provides an example of the application of the steps. Hence, though implicitly, the article explores the strengths to gain from the application of the procedure. The article supports the argument offered, among others, by Kress and van Leeuwen (2006), and Machin and Myer (2012), who argue that messages sent in a text using two different semiotic modes at the same time, verbal and visual, should be analyzed by a tool of analysis that combines the analysis of verbal and visual such as Systemic Functional Linguistics (SFL) for the analysis of verbal, and Reading Images for the analysis of the images. These, work together to unearth the messages to allow us better interpretation and understanding of the messages. The relations between verbal-visual are also discussed in this article*

**Keywords:** Multimodality, semiotic mode, text, verbal, images



## PENDAHULUAN

*Multimodality* adalah istilah yang digunakan untuk merujuk pada cara orang berkomunikasi menggunakan *modes* yang berbeda pada saat bersamaan (Kress and van Leeuwen, 1996), yang dapat didefinisikan sebagai “penggunaan beberapa *semiotic modes* dalam desain produk, atau peristiwa semiotik secara bersamaan, dan dengan cara tertentu mode-mode ini digabungkan untuk—memperkuat, melengkapi, atau berada dalam susunan tertentu” (Kress and van Leeuwen, 2001). *Multimodality* dapat juga dikatakan sebagai “istilah teknis yang bertujuan menunjukkan bahwa pemaknaan yang kita lakukan selama ini memanfaatkan beragam *semiotic*” (Iedema, 2003). Sementara Chen (2010) memaknai *multimodality* sebagai “memahami bagaimana sumber semiotik verbal dan visual dapat digunakan untuk merealisasikan jenis dan tingkatan *dialogic engagement*, keterlibatan dialogis (cetak miring dari Peneliti) dalam sebuah buku teks”. Dalam konteks analisis teks, *multimodality* menurut Peneliti dapat dipahami sebagai sebuah ‘prosedur analisis’ yang menggabungkan alat dan langkah analisis linguistik seperti misalnya *systemic functional linguistics* (SFL), atau Tata Bahasa Fungsional, dengan alat analisis untuk memahami gambar, bila teks yang dianalisis menggunakan dua mode, verbal dan gambar.

Sekarang ini, semakin banyak orang mengandalkan penggunaan mode berbeda untuk berkomunikasi. Komunikasi, dalam artikel ini dimaknai tidak saja sebagai peristiwa penyampaian pesan dari pengirim kepada penerima, tapi juga usaha penawaran makna dari pembuat pemaknaan kepada audiens, termasuk di dalamnya konsumen seperti dalam iklan sebuah produk. Sementara *mode* menurut Bezemer dan Kress (2008:171) adalah “*a socially and culturally shaped resource for making meaning*”, mode dipahami sebagai sumber yang terbentuk secara sosial dan

budaya untuk mengomunikasikan makna. Sumber ini tidak terbatas pada bahasa saja sebagai sumber yang tidak asing dan biasa digunakan, tapi juga dapat dirujuk pada gambar, bunyi, spasi/ruang yang dengannya orang menyampaikan pesan dan menawarkan makna. Sehubungan dengan ini, Kress dan van Leeuwen (2002) bahkan berargumen bahwa warna adalah juga salah satu *semiotic mode* sebab makna sebuah warna bisa berbeda dalam konteks berbeda, dan dari warna-warna lainnya. *Layout*, termasuk di dalamnya *blank space*, ‘ruang kosong’ seperti pada koran adalah contoh lain *semiotic mode* (Kress and van Leeuwen, 2006). Dalam tulisan ini, lebih jauh Peneliti melihat *semiotic mode* sebagai alat dan cara yang tersedia dan berterima dalam sebuah budaya, yang dapat digunakan untuk menyampaikan pesan yang ingin disampaikan, dan makna, representasi cara pandang, serta pemaknaan terhadap dunia yang akan ditawarkan.

Diantara mode berbeda yang digunakan orang secara bersamaan dalam sebuah teks untuk mengomunikasikan pesan mereka, mungkin gabungan verbal dan citraan atau gambar yang paling sering digunakan, dibandingkan kombinasi mode berbeda lainnya seperti misalnya verbal dan musik. Sekaitan dengan ini, teks bermakna lebih luas dari sekedar “tidak saja tulisan atau lisan...” seperti menurut Halliday (dikutip dalam Eggins, 2004), namun teks adalah “*all practices that signify*”, segala praktek yang memiliki makna, yang dapat berbentuk citraan, objek, praktek bahkan bunyi (Barker, 2000). Bersandar pada ini, setidaknya bagi Peneliti, teks dapat dirujuk pada dua hal yang berbeda. Pertama, makna teks dapat saling menggantikan dengan makna (*semiotic*) mode. Kedua, teks dapat juga dipahami sebagai ‘wadah’, situs, di mana *semiotic mode* digunakan untuk menawarkan makna. Sebagai contoh, sebuah poster secara keseluruhan dikatakan sebagai sebuah teks, yang di dalamnya terdapat

verbal dan gambar sebagai dua mode semiotik yang digunakan untuk menyampaikan pesan.

Dalam kaitannya dengan alat analisis teks multimodal, khususnya yang menggabungkan verbal dan visual, Kress dan van Leeuwen (1996, 2006) menawarkan apa yang dianggap sebagai 'tata bahasa' gambar. Mengembangkan lebih jauh pemikiran para pendahulunya, di antaranya pemikiran Arnheim (1969), mereka menggagas 'tata bahasa' gambar menggunakan perspektif bahasa. Selama ini, mereka berargumen, teks di luar bahasa seperti terutama gambar, desain, atau ilustrasi, menjadi milik kalangan tertentu yaitu para desainer atau ilustrator yang memiliki akses terhadapnya. Para desainer dan ilustrator ini, dalam pandangan Kress dan van Leeuwen, tidak memiliki cara untuk menyampaikan pengetahuan mereka dalam bahasa yang dipahami mereka yang berada di luar lingkarannya. Bersandar pada Halliday (1994, 2004) yang melihat tata bahasa tidak sebagai aturan tapi sebagai sumber untuk membuat dan menyampaikan makna, Kress dan van Leeuwen melihat bahwa imej dapat 'diperlakukan' seperti bahasa, memiliki ketiga metafungsi *ideational/logical*, *interpersonal*, dan *textual*. Bila demikian lanjut mereka, maka gambar atau imej dapat memiliki 'tata bahasa' yang dengannya orang dapat membaca imej dengan 'benar' untuk mengungkap makna yang disampaikan. 'Tata bahasa' ini kemudian dikembangkan lebih jauh misalnya oleh Machin and Myer (2012), dan O'Halloran (2008) yang menggabungkannya dengan *systemic functional linguistics*. Sementara analisis teks multimodal yang menggabungkan bahasa dengan *semiotics resources* di luar bahasa, telah banyak dilakukan misalnya oleh Steiner, dan van Leeuwen yang menganalisis musik dan bunyi, Martinec yang menganalisis gerakan dan bahasa tubuh (*gesture*), O'Toole, Pang, dan Stenglin yang masing-masing menganalisa arsitektur dan ruang (dikutip dalam O' Halloran, 2008).

'Tata bahasa' ini kemudian banyak dianggap sebagai dasar penting kerangka analisa *multimodality*, dan bersandar pada kerangka ini banyak kajian telah dilakukan misalnya, Kress dan van Leeuwen (1998) yang mengungkap ideologi dalam halaman muka koran, Royce (2002) yang mengeksplorasi sinergi antara visual dan verbal di dalam kelas, O'Halloran (2008), yang mengonstruksi makna *ideational* dalam bahasa verbal, dan visual, Koller (2008), yang meneliti fungsi warna merah muda sebagai penanda gender dan seksualitas dalam model budaya dan teks multimodal, Chen (2010) yang melihat bagaimana sumber multimodal dalam buku teks EFL di China digunakan untuk memungkinkan terciptanya keterlibatan dialogis dengan pembaca, Archer and Stent (2011) yang membahas warna, berdasarkan ide Kress dan van Leeuwen (2002, 2006), sebagai sebuah *semiotic mode*, dan cakupan sejauh mana warna dapat beroperasi sebagai sebuah mode yang berdiri sendiri.

## METODE

Kress dan van Leeuwen (1996, 2006) tidak secara eksplisit mengurutkan satu persatu langkah analisis menggunakan teorinya, tetapi mereka mengelaborasi dengan rinci poin-poin penting yang harus diperhatikan ketika melakukan analisis teks visual menggunakan *Reading Images*. Menurut keduanya, imej dapat dianggap seperti bahasa verbal, merealisasikan ketiga metafungsi bahasa. Tulisan ini merangkum semua poin penting itu dan mengurutkannya menjadi langkah analisa praktis yang dapat digunakan untuk menganalisa teks gambar. Selain terutama berdasarkan penjelasan mereka, langkah analisa dalam tulisan ini juga merujuk pada Unsworth (2001), dan Machin and Myer (2012).

Pertama, menggunakan langkah analisis ini kita harus memperlakukan imej seperti bahasa. Ini artinya, imej diyakini sebagaimana bahasa verbal

merealisasikan metafungsi *ideational*; bagaimana gambar merepresentasikan pengalaman. Dalam imej, hal ini dapat dilihat pada bagaimana misalnya objek atau *represented participants* atau item, dalam imej ‘berhubungan’ dengan objek lain. *Represented participants* secara sederhana dapat dimengerti sebagai objek yang ada dalam gambar; bisa berupa benda hidup bisa juga berupa benda tak hidup. Sementara, ‘viewer’ atau yang melihat objek dinamakan ‘interactive participant’. *Represented participant* (objek/item dalam gambar) bisa dihubungkan dengan cara apakah objek terlibat dalam proses ‘berinteraksi’ atau ‘berkoneksi’ (klasifikasi). Ini diwujudkan dalam bentuk vektor. Ketika melakukan analisis, kita harus melihat dari mana vektor berasal, dan ke mana dia bergerak. Vektor memposisikan *participant* sebagai aktor, reaktor, objek/goal, fenomena, atau pembicara. Berdasarkan vektor, hubungan antar objek dalam gambar bisa ‘transactional’, ‘non-transactional’, ‘bi-directional’, atau ‘conversion’ (lihat Kress dan van Leeuwen, 1996, Unsworth, 2001). Metafungsi kedua yang direalisasikan oleh imej adalah *interpersonal*. Ketika menganalisis imej, kita melihat secara kritis bagaimana hubungan yang tercipta dan dimiliki antara pembuat, yang melihat, dan objek yang ada dalam gambar. Dalam imej, ini direalisasikan melalui *gaze* (tatapan, dan arah tatapan), ukuran *frame* dan *shot*, serta perspektif/angle. Ketiga realisasi ini menggambarkan ‘tuntutan’, atau ‘tawaran’, jarak sosial, (intim, dekat, jauh, atau publik), kuasa serta sikap yang dimiliki oleh objek terhadap yang melihat (*viewer*) dan sebaliknya. Metafungsi ketiga yang direalisasikan imej adalah *textual*. Artinya, kita harus melihat bagaimana imej disusun dan disajikan. Seperti dalam kalimat pada bahasa verbal, bagaimana elemen dalam kalimat disusun akan memengaruhi makna kalimat secara keseluruhan. Susunan komposisi berbeda dalam imej memungkinkan makna tekstual serta nilai informasi yang berbeda

juga. Beberapa susunan komposisi yang mungkin dalam imej diantaranya *Given-New* (kanan–kiri), *Ideal-Real* (Ideal-Nyata), *Centre-Margin* (pusat-pinggir), Terpolarisasi, dan *Triptych*. Susunan komposisi ini juga memengaruhi, meski tidak selalu menentukan, alur baca (*reading path*) mereka yang melihat imej.

Selain poin-poin yang disebutkan di atas, ketika melakukan analisa pada gambar, kita juga harus mempertimbangkan *framing* dan warna yang digunakan. Unsworth (2001:109) mendefinisikan *framing* sebagai “elemen atau kumpulan elemen dalam sebuah *layout* dapat diputuskan dan dibedakan satu dengan lainnya atau dihubungkan, disambungkan bersama”. Sekaitan dengan *framing*, Goffman (dikutip dalam Rodriguez dan Dimitrova, 2011) memberikan ide dasar bahwa “konteks dan pengorganisasian pesan memengaruhi pikiran dan tindakan selanjutnya audiens terhadap pesan itu”. Dalam tulisan ini, *framing* dimaknai sebagai bagaimana elemen dalam imej ditampilkan yang dengannya, cara pandang, sikap, dan tindakan audiens terhadap apa yang ditampilkan dapat dipengaruhi. Sementara itu, warna juga dianggap memiliki makna tertentu yang umumnya dipengaruhi oleh situasi dan budaya di mana warna itu digunakan. Kress dan van Leeuwen (2002) berargumen bahwa warna adalah sebuah *semiotic mode*, sebab warna memiliki dan dapat digunakan untuk menyampaikan makna.

Kedua, setelah melihat imej dengan kerangka pandang tiga metafungsi bahasa, langkah selanjutnya dalam prosedur analisa gambar adalah penggambaran, atau identifikasi. Seringkali menggambarkan sebuah objek dipandang lebih gampang dari menganalisisnya. Menganalisa dianggap jauh lebih hebat daripada menggambarkan atau mendeskripsikan, salah satunya karena pandangan bahwa dalam menganalisa kita menggunakan nalar kritis kita untuk menemukan hubungan yang satu dengan yang lain.

Padahal, untuk membuat sebuah gambaran yang baik dibutuhkan ketelitian; dalam menggambarkan imej khususnya kita tidak saja menggambarkan apa yang ada namun seringkali juga harus mengidentifikasi apa yang tak ada, namun terasa penting, dalam imej yang kita lihat. Haruslah dipahami bahwa tidak akan ada analisa yang baik tanpa penggambaran yang rinci. Bagi Machin dan Myer (2012) mendeskripsikan apa yang ada pada imej 'memfasilitasi analisa yang baik'. Disarankan agar kita menggambarkan secara rinci setiap elemen utama dalam imej termasuk bagaimana elemen ini ditampilkan. Setelah penggambaran atau

identifikasi, analisa kritis atau signifikansi kemudian dilakukan. Dalam konteks ini, penguasaan dan pemahaman makna konotasi dan denotasi elemen dalam konteks gambar yang sedang dianalisa sangat membantu.

### PENERAPAN DAN HASIL ANALISIS

Berikut ini contoh sederhana bagaimana langkah analisa yang dijelaskan di atas dapat diterapkan. Analisa dilakukan pada dua buah gambar (1 dan 2) yang berperan sebagai ilustrasi, yang diambil dari buku bahasa Inggris untuk sekolah dasar yang diterbitkan salah satu penerbit di Bandung.



Gambar 1. (Penggambaran/identifikasi) Gambar ini adalah gambar dapur.

Pada gambar di atas ada dua *participants*; seorang anak yang berkata "This is the kitchen. My mom is cooking", dengan kedua tangan yang terentang memakai baju warna biru dengan kerah warna oranye dan celana panjang coklat. Posisi anak di sebelah kanan gambar. Tatapannya meskipun tidak terlalu jelas, menuju pada *participant* lainnya, seorang ibu yang sedang memasak. Ibu, di sebelah kiri gambar, berdiri dengan posisi sedikit menyamping dan tatapannya tidak ditujukan pada anak tetapi mengarah pada apa yang ada di depannya. Di depan samping ibu adalah wajan dan katel. Wajan, katel, meja, dan benda lainnya menjadi seting di mana proses berlangsung. (Analisis/signifikansi) Vektor dalam gambar ini direalisasikan

melalui tatapan mata, bergerak dari anak menuju ibu. Dengan demikian anak menjadi 'reacter' (*reactor*) sementara ibu menjadi fenomena; seseorang yang kegiatannya diceritakan. Menurut Unsworth (2001) bersandar pada Kress dan van Leeuwen (1996), ketika vektor dibentuk oleh garis/tatapan mata salah satu atau lebih 'participant' sehingga mereka melihat pada sesuatu, maka proses yang terjadi dipandang lebih sebagai sebuah reaksi daripada aksi, dan 'participants' disebut 'reacter' bukan 'actor'.

Objek dalam gambar ini digambarkan dalam ukuran (*size*) yang kecil, dan menggunakan 'medium close shot', menempatkan kita sebagai 'viewer' sedikit di atas gambar; ketika kita melihat

gambar kita seperti sedikit melihat ke bawah. Ini dapat dimaknai bahwa sebagai 'viewer' kita memiliki kuasa lebih daripada 'represented participant' yang ada dalam gambar. 'Medium close shot' menjadikan kita hanya memiliki hubungan sosial bukan sangat dekat dengan 'represented participant'. Kita mengenal mereka seperti kita mengenal orang secara umum saja; mereka bukan bagian orang-orang terdekat kita. Lebih jauh ini bermakna, 'represented participants' mewakili gambaran umum ibu dan kegiatan apa yang umum dilekatkan pada peran domestik ibu dalam praktek sosial masyarakat kita.

Tatapan kedua *participant* tidak ditujukan pada kita yang melihat tapi ke arah lain sehingga tidak ada kontak antara mereka dengan kita. Ketika tidak ada kontak tatapan mata antara 'represented participants' dengan 'viewer' maka 'viewer' ditempatkan sebagai pengamat (*observer*); mereka tidak 'menuntut' tetapi 'menawarkan' pada kita. Sebagai pengamat (*observer*), kita ditawarkan untuk mengobservasi apa yang terjadi di dapur. Ukuran 'participant' dalam gambar kecil begitu juga dengan *frame*-nya. Ini membuat kita sebagai 'interactive participant' memiliki kuasa yang lebih besar terhadap 'represented participant'.

Dari sudut pandang *compositional layout*, posisi ibu ada di kiri gambar, sementara anak berada di sebelah kanan gambar. Menurut Kress dan van Leeuwen apa yang ditempatkan di kiri gambar adalah 'given' sementara apa yang di sebelah kanan adalah 'new'. Pada gambar di atas, ibu dan apa yang dilakukannya adalah 'given' sementara anak adalah 'new'. Dalam konteks ini, perlu dijelaskan bahwa komposisi 'given-new' berlaku untuk budaya di mana alur membaca bergerak dari kiri ke kanan, sementara untuk budaya yang menggunakan alur baca kanan ke kiri atau atas ke bawah seperti budaya Arab dan China, komposisi 'given-new' tidak dapat digunakan. Budaya seperti itu menggunakan istilah komposisi *layout*

yang berbeda. Bentuk komposisi ini telah mendapat kritikan karena dianggap di antaranya, menggunakan generalisasi berlebihan atas semua budaya. Namun demikian, bagi budaya baca tulis yang bergerak dari kiri ke kanan seperti Indonesia. Pada gambar yang dianalisis, 'given' maknanya adalah ibu dan terutama yang dilakukan ibu adalah sesuatu yang telah biasa, wajar, dan sudah seharusnya. Sementara yang baru pun lebih penting adalah apa yang dilakukan anak.

Biru sebagai warna baju anak termasuk warna lembut dan biasanya dianggap menggambarkan ketenangan. Ini karena biru sering dikaitkan dengan warna laut atau gunung. Warna baju ibu, oranye yang 'dibungkus' coklat. Oranye termasuk warna 'hangat', biasanya dianggap di antaranya menggambarkan semangat. Sementara warna coklat, warna yang 'membungkus' warna oranye, adalah warna lembut. Dapat dimaknai bahwa ibu bersemangat, senang melakukan apa yang dianggap 'seharusnya' dilakukan oleh seorang ibu/istri.

Analisis verbal menunjukkan bahwa kalimat yang diucapkan anak mengandung proses *relational* (*This is the kitchen*) dan material (*is cooking*). Klausa relasional berfungsi untuk memberi ciri dan untuk mengidentifikasi. Sementara klausa material digunakan untuk menunjukkan, 'doing' or 'happening', melakukan sesuatu, atau kejadian yang sedang berlangsung (Halliday, dan Matthiesen, 2004). Dalam konteks gambar 1, kalimat yang diucapkan anak mengidentifikasi ruangan sebagai dapur, salah satu cirinya adalah adanya alat memasak dan seseorang, dalam hal ini, ibu menjadi aktor yang sedang melakukan pekerjaan yang sudah 'seharusnya' dilakukan, yaitu memasak di dapur.

Di bawah ini adalah contoh analisis untuk gambar (2) yang muncul bersamaan dengan gambar (1) dalam buku yang sama.







Gambar 2. (Penggambaran/identifikasi)

Gambar ini adalah rangkaian gambar berikutnya, yang muncul setelah gambar 1 di atas. Pada gambar ini ada dua 'participant'; anak, masih menggunakan pakaian yang sama dengan yang dipakai di gambar 1, berkata "This is my living room. My dad is reading a newspaper" dengan satu tangan menunjuk ke arah 'participant' lainnya, seorang ayah yang memakai celana panjang warna coklat, duduk di atas sofa warna hijau muda, memakai baju warna abu-abu muda, membaca koran. Posisi anak ada di sebelah kiri gambar sementara ayah ada di sebelah kanan gambar. Tatapan mata anak seharusnya tertuju pada ayah, tetapi dalam gambar ini, tatapan mata anak tampak sedikit tidak searah dengan arah telunjuk yang mengarah pada ayah. Seting pada gambar menunjukkan bahwa ini adalah gambar ruang tengah dari rumah yang sama dengan gambar dapur di atas. (Analisis/Signifikansi) Vektor dalam gambar direalisasikan melalui, terutama, oleh telunjuk anak. Arah tatapan anak sepertinya juga diarahkan kepada ayah, tetapi seperti tampak dalam gambar, ini tidak terealisasikan dengan bagus; arah tatapan anak sedikit tertuju pada apa yang ada di depan, sementara ayah berada di sisi jauh sebelah kanan anak. Arah vektor memposisikan anak sebagai 'actor', yang menceritakan, sementara ayah adalah 'goal' yang

kegiatannya diceritakan. Tatapan kedua 'participant' tidak diarahkan pada kita sebagai 'interactive participant'. Mereka tidak 'menuntut' kita, mereka 'menawarkan' sesuatu pada kita. Dalam konteks ini, anak menawarkan pada kita untuk bersama mengobservasi apa yang ayah lakukan. Ukuran gambar dan 'participant' dalam gambar kecil sehingga kita, sebagai 'represented participant' memiliki kuasa yang lebih besar terhadap mereka, dibanding mereka terhadap kita. Posisi anak yang ada di kiri adalah 'given'. Bila dihubungkan dengan gambar sebelumnya, anak pada gambar ini bukan lagi 'sesuatu yang baru', dia telah menjadi sesuatu yang sudah dikenali. Dengan demikian dia sudah menjadi 'biasa'. Sementara, ayah adalah 'new'; ayah dan apa yang ayah lakukan menjadi sesuatu yang baru, yang penting dalam gambar, dan harus menjadi fokus perhatian. Ini sesuai dengan apa yang disampaikan rangkaian gambar; gambar 2, menjadi sesuatu yang baru setelah gambar 1 dilihat oleh pembaca.

Analisa verbal menunjukkan dua proses yang digunakan, relational (*this is my living room*), dan material (*my dad is reading a newspaper*). Klausa relasional mengidentifikasi ruangan, menunjukkan apa yang menjadi ciri ruangan. Ciri ruang tengah adalah adanya meja, sofa, dan hiasan dinding. Ruang tengah adalah

ruang yang digunakan salah satunya untuk (ayah) membaca. Klausula material digunakan untuk menunjukkan bahwa ayah adalah seseorang yang melakukan sesuatu yang 'sewajarnya' dilakukan ayah, yaitu membaca koran di ruangan tengah.

Dari sudut pandang hubungan verbal dan visual dalam kedua gambar di atas, teks verbal dalam bentuk 'speech balloon' memperluas makna yang disampaikan gambar dan sebaliknya, gambar memperluas makna yang disampaikan apa yang diucapkan dalam 'speech balloon'. Ini didasarkan pada pendapat Barthes (dikutip dalam Kress dan van Leeuwen, 2006:18) "... *the verbal text extends the meaning of the image, or vice versa, as in the case, for example, the speech balloon in comic strip ...*". Teks verbal memperluas makna gambar, atau sebaliknya, seperti dalam kasus, misalnya, balon percakapan dalam komik. Pada kedua gambar yang dianalisis tidak ada pertentangan makna antara teks verbal dan visual. Mereka bekerja menyampaikan pesan yang sama.

Secara keseluruhan, pesan yang disampaikan kedua ilustrasi dari buku teks bahasa Inggris untuk SD ini adalah, ibu dan ayah memiliki ruangan yang menjadi penanda identitas gender sekaligus menunjukkan tugas domestik yang menjadi ciri yang melekat di dalamnya. Ibu bertugas memasak, menyiapkan makanan untuk disantap keluarganya. Pada saat yang bersamaan Ayah yang adalah kepala rumah tangga, dapat menikmati waktu untuk membaca koran di ruang tengah. Ini karena tugas Ayah adalah memenuhi kebutuhan keuangan keluarga.

Bila kita meyakini bahwa teks, apapun bentuknya, adalah 'recontextualization of social practice', kontekstualisasi ulang dari praktek sosial (van Leeuwen, 2008, 2009), maka kita harus melihat bahwa apa yang ada dalam gambar 1 dan 2 dari buku teks yang dianalisis, bukan sekedar ilustrasi semata tapi merupakan cerminan, gambaran

(ulang) dari apa yang menjadi praktek sosial yang berlaku di masyarakat di mana buku teks berada. Meskipun tentu ada alternatif berbeda yang mungkin dari praktek sosial yang terjadi, kenyataan bahwa apa yang ada dalam gambar yang dipilih menunjukkan bahwa ini adalah keyakinan yang dipandang sebagai sebuah kebenaran. Dengan kata lain, ini adalah ideologi buku teks ini.

Dengan demikian, meskipun tidak ada yang salah dengan kenyataan bahwa seorang ibu memasak, melakukan salah satu tugasnya sebagai ibu rumah tangga, pesan dalam buku teks ini dapat dipahami sebagai 'reproduksi' dan 'pemeliharaan' ideologi patriarki, yang menempatkan perempuan sebagai pengurus pekerjaan rumah tangga sementara laki-laki adalah pencari uang dan tidak harus melakukan urusan dapur.

## **HUBUNGAN YANG VERBAL DENGAN YANG VISUAL**

Imej, atau gambar, memiliki beragam hubungan yang mungkin dengan verbal, kata-kata. Bagi Barthes (dikutip dalam Kress dan van Leeuwen, 2006: 18) verbal menambah/memperluas (*extend*) makna imej dan sebaliknya. Atau, verbal menjelaskan (*elaborate*) imej dan sebaliknya. Kress dan van Leeuwen lebih jauh mengatakan bahwa bagi Barthes, makna imej khususnya, dan *semiotic mode* lainnya, selalu dihubungkan dengan, dan tergantung pada makna teks verbal. Sementara bagi Kress dan van Leeuwen, komponen visual teks, imej, adalah sebuah pesan yang disusun dan diatur secara mandiri, berhubungan dengan teks verbal tapi tidak *tergantung* (garis miring dari Peneliti) padanya. Chen (2010) menggunakan empat kategori *heteroglossic dimension* yang diajukan Martin dan White (dikutip dalam Chen, 2010), menemukan bahwa 'suara' atau pesan yang disampaikan yang visual bisa bertentangan (*disclaim*) dengan 'suara' atau pesan yang disampaikan verbal. Padahal, teks verbal dan visual ini muncul secara bersamaan dalam sebuah

kesempatan. Hubungan ini juga ditemukan oleh Bednarek and Caple (2012) yang meneliti foto jurnalisisme dalam koran cetak dan daring.

Royce (2002) melihat bahwa visual dan verbal yang digunakan sebagai mode untuk menyampaikan pesan dalam sebuah teks memiliki ‘intersemiotic relationship’, hubungan antar mode semiotik yang beragam. Hubungan antara keduanya bisa merupakan ‘intersemiotic repetition’, *intersemiotic synonymy* (similarity relations), *intersemiotic antonymy* (opposition relations), *intersemiotic hyponymy* (class-subclass relations), *intersemiotic meronymy* (part-whole relations), *intersemiotic collocation* (expectancy relations).

Lebih jauh, untuk menjelaskan hubungan teks verbal dan visual, Martinec dan Salway (2005) menawarkan sebuah sistem untuk melihat hubungan antara visual dan verbal dalam sebuah teks multimodal. Pendekatan mereka berdasarkan hubungan status dan logis-semantik (*logico-semantics*) antara visual dan verbal. Ini berbeda dengan sistem yang dikembangkan oleh Liu dan O’Halloran (2009) yang lebih berdasar pada ‘discourse relation’ antara teks verbal dan visual dalam analisis mereka pada ‘cohesive devices’ antara teks verbal dan visual.

## SIMPULAN

Tulisan sederhana ini telah menjelaskan makna *multimodality*, memberikan panduan langkah analisis gambar, *semiotic mode lain*, yang sering digunakan bersama dengan mode verbal, menggunakan teori *reading images*. Tulisan ini juga telah memberikan contoh penerapan ‘prosedur’ analisis, yang melaluinya secara implisit ditawarkan manfaat yang dapat diambil dari langkah analisis ini.

Dalam dunia yang kini diserbu teks multimodal, kita tidak lagi bisa hanya mengandalkan penguasaan dan pemahaman kita pada satu *mode* saja untuk memahami pesan yang

disampaikan teks secara menyeluruh dan lebih ‘benar’. Kita harus menguasai cara membaca gambar, sesuatu yang selama ini dianggap tak cukup diwakili oleh seribu kata. Paling tidak, dengan menguasai dan memahami alat analisis ini kita bisa mendeskripsikan gambar, ilustrasi, lukisan dan desain dengan lebih baik tanpa harus bergantung pada apa yang disampaikan oleh para kurator lukisan, pembuat ilustrasi, atau disainer. Di samping itu, kemampuan secara fasih dan berdasar untuk menggambarkan sebuah fenomena termasuk di dalamnya gambar, imej, selain memberikan kontribusi nyata dan bermakna bagi perbaikan dunia, adalah apa yang membedakan akademisi—yang tercerahkan—dengan mereka yang bukan akademisi.

Mungkin di masa datang yang tak akan lama lagi, apa yang membuat kita mampu menjalankan tugas kita sehari-hari dengan baik, memampukan kita menghadapi rutinitas kehidupan dengan lebih bermakna, menyelamatkan kita dari bahaya dan memungkinkan kita menjadi orang yang berkontribusi bagi dunia, bukan lagi pemahaman kita pada arti kata-kata saja, tapi pemahaman kita pada apa arti kata dan gambar, imej, dengan pemahaman kita pada imej memainkan peranan yang lebih besar dan menentukan.

## UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada Prof. Theo van Leeuwen, Dekan Fakultas Komunikasi University of Technology (UTS) Sydney Australia, yang telah memberi kesempatan pada Peneliti untuk menimba ilmu tentang *Multimodality* dan (*Critical*) *Discourse Analysis*. Terima kasih juga untuk John Knox, Ph.D dari Macquary University Sydney Australia, yang telah mengizinkan Peneliti untuk mengikuti kuliah tentang *Multimodality*. Kemudian ucapan terima kasih juga untuk Helen Caple, Ph.D, dari University of New South Wales Sydney Australia, yang bersedia menjawab

banyak pertanyaan selama Peneliti melakukan *Visiting Scholar* di UTS, dari 7 September sampai 27 november 2012. Diskusi yang bernash dengan mereka menjadi salah satu pendorong Peneliti membuat artikel ini.

#### PUSTAKA RUJUKAN

- Archer, Arlene & Stent, Stacey. 2011. Red Socks and Purple Rain: The Political Uses of Colour in Late Apartheid South Africa. *Visual Communication*, 10: 115.
- Arnheim, R. 1969. *Visual Thinking*. Berkeley and Los Angeles: University of California Press.
- Barker, C. 2000. *Cultural Studies: Theory and Practice*. London : SAGE Publications. Ltd.
- Bednarek, M. dan Caple, H. 2012. *News Discourse*. London and New York: Continuum .
- Bezemer, J. & Kress, G. 2008. Writing in Multimodal Texts: A Social Semiotic Account of Designs for Learning. *Written Communication*, 25 : 166.
- Chen, Y. 2010. Exploring Dialogic Engagement with Readers in Multimodal EFL Textbooks in China. *Visual Communication*.
- Egins, S. 2004. *An introduction to Systemic Functional Linguistics*, 2<sup>nd</sup> ed. London: Continuum International Publishing Group.
- Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*, second edition. New York: Routledge, Chapman and Hall, Inc.
- Halliday, M.A.K. dan Matthiessen, Christian, M.I.M. 2004. *An Introduction to Functionnal Grammar*, 3<sup>rd</sup> edition. Great Britain: Arnold.
- Iedema, R. 2003. Multimodality, Resemiotization: Extending the Analysis of Discourse as Multi-semiotic Practice. *Visual Communication*, 1-30.
- Koller, V. 2008. 'Not just A Colour': Pink as A Gender and Sexuality Marker in Visual Communication. *Visual Communication*, 7: 395.
- Kress, G. & Van Leeuwen, T. 2006. *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kress, G. & Van Leeuwen, T. 1996. Front Pages: (The Critical) Analysis of Newspaper Layout. In *Bell, Allan. and Garret, Peter (Eds), Approaches to Media Discourse*. Oxford: Blackwell.
- Kress, G. & Van Leeuwen, T. 1998. Front Pages: (The Critical) Analysis of Newspaper Layout. In *Bell, Allan. and Garret, Peter (Eds), Approaches to Media Discourse*. Oxford: Blackwell.
- Kress, G. & Van Leeuwen, T. 2001. *Multimodal Discourse The Modes and Media of Contemporaray Communication*. Great Britain: Arnold.
- Kress, G. & Van Leeuwen, T. 2002. Colour as A Semiotic Mode: Notes for A Grammar of Colour. *Visual Communication*, 1-27.
- Liu, Yu & O'Halloran, Kay L. 2009. Intersemiotic Texture: Analyzing Cohesive Devices between Language and Images. *Social Semiotic*, 367-388.
- Machin, D & Myer, A. 2012. *How to Do Critical Discourse Analysis*. London: SAGE Publication Ltd.
- Martinec, Radan & Salway, Andrew. 2005. A System for Image-Text Relation in New (and Old) Media. *Visual Communication*, 4:337
- O'Halloran, K. L. 2008. Systemic Functional-Multimodal Discourse Analysis (SF-MDA): Constructing Ideational Meaning Using Language and Visual Imagery. *Visual Communication* , 443-473.
- Rodriguez, L & Dimitrova, Daniela. 2011. The Levels of Visual Framing. *Visual Literacy*, 48-65.

- Royce, T. 2002. Multimodality in the TESOL Classroom: Exploring Visual-Verbal Synergy. *Tesol Quarterly*, Vol. 36 No. 2 Summer.
- Unsworth, L. 2001. Describing Visual Literacy. In *Teaching Multiliteracy Across Curriculum: Changing context of text and image in classroom* (pp. 71-112). Buckingham; Philadelphia: Open University Press.
- Van Leeuwen, T. 2008. *Discourse and Practice. New Tools for Critical Discourse Analysis*. New York: Oxford University Press.
- Van Leeuwen, T. 2009. Discourse as the Recontextualization of Social Practice: A Guide. Dalam Ruth Wodak dan Michael Meyer (Eds), *Methods of Critical Discourse Analysis, 2ed*. London: SAGE Publications Inc.