
It is made available according to the terms of agreement between the author and the journal, and in accordance with UCT's open access policy available: http://www.openuct.uct.ac.za/sites/default/files/UCTOpenAccessPolicy.pdf, for the purposes of research, teaching and private study.
Biographical Statement

Dr Arlene Archer is the co-ordinator of the Writing Centre at the University of Cape Town, South Africa. She teaches in Applied Language Studies, Higher Education Studies, Film and Media, as well as in an Engineering Foundation programme. Her research focuses on the possibilities multimodal pedagogies offer for enabling student access to Higher Education. She has published in journals such as Language and Education, Teaching in Higher Education, English in Education, English for Specific Purposes.
Multimodal texts in Higher Education and the implications for writing pedagogy

Abstract
Although studies on writing pedagogy and academic literacies have examined changing genres in tertiary education, there has not necessarily been an emphasis on how a range of modes and media have influenced texts in various disciplines. This paper explores the influence and incorporation of the visual into student texts in Higher Education, looking at the semiotic weighting of modes, conventions and functions of images, visual / verbal linkages and visual composition. These aspects of multimodal texts have implications for the ways in which we teach ‘academic literacies’ practices.

Key words
Academic literacies, academic discourse, multimodality, visual/verbal linkages, modal specialization, modal weighting

Permission has been obtained from the students to use abstracts from their texts.
Multimodal texts in Higher Education and the implications for writing pedagogy

Academic texts of many kinds rely on the co-presence of writing and graphic material. Images are becoming increasingly important as carriers of meaning in a broad range of academic disciplines such as art, art history, architecture, film studies, cultural studies, media studies, communication science, historical studies, literary studies, anthropology, sociology, philosophy, psychology and education (Van Leeuwen 2005: 11 – 12). Although there is an increasing incorporation of images into student writing in Higher Education, there has not been much exploration into what academic discourse looks like across visual and verbal modes. Student-produced texts in Higher Education utilize images in a range of ways. Images are used in conjunction with the written mode to provide context, illustrate a point, make an argument, furnish evidence, organize data. There are also a range of examples of academic genres that require the use of numerical graphical representations, such as needs analyses, impact assessments, cost benefit analyses. This article attempts to make explicit some of the ways that academic discourse is constructed in a multimodal environment in order to enable student access to this specialized discourse.

The teaching of academic literacy practices in Higher Education is inextricably linked to student access, which includes both retention and throughput. If the requirements of academic assignments are increasingly multimodal, we need to be equipped to help students with these new and emerging texts. This is not a ‘technical’ issue. Rather, multimodal texts in the curriculum raise
fundamental questions about power and access. For instance, multimodality in the humanities requires students to engage with four ‘kinds’ of language: the English language system, academic discourse, mode-specific language associated with the analysis of the visual and a metalanguage of critical analysis (Thesen 2001). This can prove to be extremely complex, and how to facilitate access to these ‘languages’ needs to be carefully considered.

[Put in def of multimodality – see powerpoint]

The approach to texts taken by this study is multimodal social semiotics. The assumption underpinning this approach is that meaning is made through the selection and configuration of modes in texts and through the interests of the sign-maker in a particular context (Jewitt 2009: 15). As coordinator of an academic writing centre at a South African university, I became interested in the ways in which multimodal student texts are constructed and the extent to which students internalize the often unconscious practices of the discipline. Making these practices explicit and visible is becoming more and more crucial in the teaching of writing, both at the centre and in stand alone courses in the disciplines. In this paper, I look at a range of student texts from different disciplines and at different levels: a first year Engineering foundation course, a Humanities foundation course, linguistics and architecture. In these contexts, images are used to greater or lesser degrees and with different functions. In looking at these student texts, I begin by investigating how different modes are used to realize different functions in a student produced poster in engineering. I then look at the challenges of writing a reflection on a visual
production, the complexities of visual / verbal linkages, and end by exploring a written analysis of a visual text in a Humanities course.

**Modal weighting of diverse functions**

Each mode in a multimodal text realizes different functions or communicative work. In other words, different aspects of meaning are carried in different ways by each of the modes. As an example of this, I will look at a poster produced in a communication course in engineering. The course focuses on developing academic literacy practices to first year engineering students at a South African university. It is about rural development and the students have to research a particular area and make recommendations for ways of improving the infrastructure of a rural village in the form of a poster. These posters raised a number of interesting questions. They made me wonder whether different modes and genres enable different competencies, social relations, and degrees of affective involvement. It seemed to me that causal relationships between technology and society were expressed to a far greater extent in the poster genre than in students' written reports. For instance, the posters could blend pathos (through the display of a burn victim in an informal settlement) and technological detail, implying that technological advancement can address social ills. I became interested in focusing on students' interest and motivations for the uses of different modes to realize different functions. The poster I will look at here, the Ingogo Village poster, consists of a set of colour-coordinated 'boxes', each describing one aspect of development: water supply, community planting schemes, hydro power, housing and soil erosion (see figure 1). The images, comprising photographs and diagrams, are
strongly related to aspects of the written text. Although the poster is tightly framed with strong vertical and horizontal lines, there is no ‘centre’ to the representation and there is no clearly designated reading path.

Figure 1: Engineering students’ poster representing proposed developments in a rural area

There is a paradigmatic logic at work in the poster, where meaning is not derived from sequence, but the relation of parts to the whole is exemplified. The whole is ‘rural development’ and the ‘parts’ are the different infrastructural aspects of the village (housing, water supply and sanitation, the environment, hydro power). This conceptual framework is reinforced by the use of colour and the compositional design, where colours are used for the classification of the parts: pink for water supply, red for the introduction and conclusions, blue for power, yellow for housing and orange for the environment. The conceptual
framework is also realized typographically by consistency in the headings of the units (in terms of size, case, font) indicating equality rather than a hierarchical relationship between units. So, although the written mode dominates over the visual in terms of conveying information and space allocation on the poster, the logic of the design is a spatial logic rather than a linear one.

In terms of modal weighting, the images chosen seem to realize the affective dimensions of the representation, rather than the scientific. The photograph to illustrate the community planting scheme depicts a boy holding up his produce. His gaze is that of direct address to the audience and he provides a ‘human face’ to the developments proposed for the village. Also, a sense of utopianism and possibility seems to be realized in the visual mode, whereas the written mode is grimly pragmatic (emphasizing the constraints of the proposals put forward and the reality of crime in the area). The pictures show what ‘could be’, the possibilities for development in the village, rather than the existing situation. The table depicts the nominal quantities of daily water supply that should be available for consumption and agriculture rather than the current supply. The photographs of the windmill and indeterminate open land, depict wide, open horizons. The community planting scheme photograph shows a small ploughed field which is ordered and productive. The fruits of production are represented in the photograph below the ploughed field, where the smiling child holds up home grown produce. Although the housing picture represents ‘house’ as a decontextualized type, an abstract notion of an architect’s plan of a house, the sweeping arc of light over the house is
reminiscent of a rainbow or a halo. The diagram representing hydro-power shows how power is generated in times of low and peak demand. This is an ordered, schematic diagram with the reassurances of operational systems in place. In general, the visuals represent systems of order and stability and hope, which is somewhat different to the bleak view represented through the written mode.

By looking at this particular student poster, I have shown the ways in which different functions can be realized through different modes in a multimodal text. Sometimes these functions complement each other, but they can also contradict each other. This raises challenges for assessment. Where different modes realize different functions in a complementary way, overall coherence is achieved. However, where there is a clear disjuncture and opposition between the functions of the modes, coherence is compromised. This kind of multimodal text can achieve different kinds of insight to, say, a written report, and these affordances need to be considered in an assessment rubric.

**Written reflection on visual productions**

The second type of multimodal student-produced text I would like to look at is a written reflection which often accompanies predominantly visual work. The example is taken from a first year assignment that requires architecture students to do a visual comparison using photoshop. Architects are often required to illustrate how spaces and facades have changed or will change in function over time. The example that students have to work with is the gentrification of an inner city area, showing ‘before’ and ‘after’ photographs
from 1958 and 2003. Using photoshop, they produce a synthesis of the two images showing what they consider to be the important changes to the street in a single image. Significant changes are conveyed visually through displaying different layers, by changing the order of layers or changing the opacity of layers. As is common with other predominantly visual genres in academia, they have to provide a short written explanation on how they conveyed change in the new image. This includes a reflection on how colour has been used to communicate change, as well as the use of opacity.

Besides these reflections on representational choices, the students need to think about larger social concerns such as how changing the way people use the pavements and the road impacts on spaces in the city. Look at the following as an example of this type of reflection:

*I have highlighted the changes I thought were particularly important, in terms of the uses of the spaces. Such as the trees which now line the street, and change the spatial arrangement. And also in that the street is not commercialised as compared to being deserted … Movement through the city, either as a pedestrian or through a vehicle, is essential in the nature of space in a city. By changing the movement through space you change the way the citizens’ senses react to the environment (narrower roads = more intimate experience), thus you change the whole experience and by this means you can manipulate or guide the experience of the space to a certain extent.*

Also, the hybrid genre of ‘reflection’, or the genre of ‘comment on …’, needs to be considered. It is clear from the above that ‘reflection’ has connotations of being more ‘personal’. This is demonstrated in the conversational style of writing which assumes an audience ‘in the know’, and at times is more akin to stream of consciousness or thinking aloud. The language of these reflections is interesting. Here the writer uses sentence fragments and run-on sentences (with three conjunctions), which are more akin to talking than to formal
academic writing. The writer also uses a visual marker, the ‘is equals to’ sign, rather than words. However, the voice of the discipline is present, drawing on technical conventions, such as the use of space and movement through space. The writing is not just reflecting on an ‘inner creative process’. The author is constructed as a participant inside the text, rather than as an outsider describing a phenomenon, which is necessary for a critical stance.

The requirement that visual texts be accompanied by some kind of written reflection in Higher Education could be a vestige of conventions requiring a written component to the assessment task. Perhaps this demonstrates the bias towards the written in this environment, but it could also point towards the affordances of the written mode for critical reflection. In designing assessment tasks, careful attention needs to be given to the relative function of the production and the written reflection, and the assessment weighted accordingly. This written reflection can be a tricky genre to produce and often requires the use of a complex metalanguage to describe and analyze the visuals. If no metalanguage is provided, students battle with vague generalities rather than insightful analyses, because they do not have the tools available to write about the images. This metalanguage can be appropriate for the level of the students and the particular context. For instance, first year architecture students may require a more explicit way of talking about images than first year engineering students.
Conventions and functions of images

In multimodal texts, it is important to note the conventions surrounding the type of image, the function of the image, and the choice of a particular image.

In terms of conventions surrounding a particular type of image, there are a number of ways of presenting summative numeric information, numeric ranges and percentages. For instance, bar charts are used to compare quantities, pie charts are used to show proportions of the whole, line graphs show quantities over time, scatter plots indicate density and frequency, timelines represent history. There are also conventions for representing relationships such as venn diagrams where circles are used to show groups and how they overlap, tree diagrams (as in family trees), linear diagrams (as in subway maps or electrical circuits). Architects and engineers use drawing conventions that present complex three-dimensional relationships in two dimensions, and hidden details can be revealed through cross-sections, cutaways or schematic drawings. Prediction can be indicated in images in a number of ways, for instance, the use of dotted lines in economics models. The use of diagrams in sequence to explain a process is common across disciplines.

In terms of the function an image is performing in a text, the image can serve as an illustration, it can be used as evidence in an argument, it can be a part of an argument (the proposition, for instance), or it can constitute ‘restatement’ in a different mode. In Prince and Author (2008), we argue that numerical graphical representations in texts often function rhetorically. These types of representations tend to have high modality or truth value in academic
texts as the assumptions underlying the numbers are generally hidden and numerical representations are often regarded as more factual and objective than other kinds of evidence.

The type of image chosen for a particular purpose is important. A photograph may be better suited to emphasize a building in its surrounding context, whereas a more abstract drawing of a plan of a building could emphasize certain structural aspects better than a photograph could. In a scientific coding orientation, a cross-section diagram, perspective drawing or simplified two-dimensional drawing may be more appropriate than a more realistic drawing or photograph. As a part of writing pedagogy, it is important to create awareness about the importance of choice of image for a particular purpose and audience, the conventions and functions of the image in a particular context. The ways the image is incorporated into the text is also important and for this reason it is worth exploring the visual/verbal relations.

**Visual/verbal relations**

Many multimodality theorists have thought about visual/verbal relations and attempted to systematically describe resources for meaning-making at the intersection of image and writing (Martinec and Salway 2005, Unsworth 2006, Royce 2002). Most of these analyses draw on a social semiotic metafunctional view of communication, specifically using concepts from systemic functional linguistics (Halliday and Hasan 1989). Although they classify and describe the visual-verbal relations differently, there are some underlying trends that emerge: similarity relations, opposition relations,
complementary relations, and whole-part paradigmatic relations. The relations between image and writing are thus classified in terms of dependency relations, namely the meaning realized by image and writing could be equal or unequal, contradictory or complementary. In terms of similarity relations, one could look at how one mode exemplifies the other. In complementary relationships, what is represented in images and what is represented in writing may be different, but complementary. One problem in looking at visual/verbal relations in this way is that the focus falls on the relation between image and caption, rather than the relation between image, caption and the larger text. Here Van Leeuwen’s (2005) notion of cohesion or information linking in multimodal compositions could be helpful.

Look at figure 2 as an example of visual/verbal linkages in a predominantly written text. This is taken from a PhD thesis looking at the linguistic features of ‘tsotsitaal’ in South Africa where ‘tsotsitaal’ refers to the language of a particular group of urban youths, often associated with criminal activities.

The image serves a contextualizing function in the PhD thesis – to locate this particular language, ‘tsotsitaal’, in its socio-cultural context. It is placed at the beginning of the thesis in a section entitled ‘History of Tsotsitaal’ and is used to flesh out the concept of a ‘tsotsi’. The image is overtly emotive (representing fury and confrontation between the represented participants) in a way that the caption and surrounding writing is not (although, a sense of the injustices of the Apartheid system does pervade both). Perhaps this is to do with the functional specialization of the visual mode.
1.5.1 History of Tsotsitaal
Tsotsis
Tsotsitaal was first observed in Johannesburg in the 1940s (Glaser 2000: 50). The second syllable, taal, means ‘language’ in Afrikaans, and the full word Tsotsitaal simply means ‘Tsotsi language’. A well grounded account for the emergence of the term tsotsi is that of Clive Glaser in his book Bo-Tsotsi (2000), who shows that it was first coined around 1943–44 in reference to a particular group of urban youths (Glaser 2000: 50). Tsotsis were both a result of, and a justification for, apartheid policies of segregation (photograph by Ernest Cole 1967).

The image is able to convey the ‘body’ of the tsotsi – the clothes, gestures, stance, gaze – in a way that may be more difficult in writing. The emotive image and less emotive writing of the caption echoes the modal distribution of functions seen in the engineering poster analyzed earlier.
The double caption here is interesting and unusual as each caption is doing different things, and together they make up a new whole. The caption immediately under the image is an example of elaboration, and more specifically ‘specification’ or ‘anchorage’ in Barthes’ (1977) terms. The words focus the reading of the image (onto some of the social and contextual circumstances) and constrain the number of possible readings one could generate. The caption used by the writer of the PhD is an example of extension, the words are complementing or enhancing the image, and the caption positions the image as part of a larger argument. As does the third caption, namely my own, which certainly does not exemplify or complement the image in any way, nor does it explain or specify anything in the image. The relationship is not quite oppositional either. It is more of a meta-level commentary on both image and captions, and draws attention to the relationship between these in order to exemplify aspects of my argument. The three captions are characterized by increasing authorial distance from the represented subject matter of the image. From the above analysis, it is clear that understanding how language and images interact to create meaning is crucial for reconceptualizing writing pedagogy from a multimodal perspective.

**Written texts that analyze and discuss visuals**

There are many examples of assignments that require students to describe, analyze or discuss visuals in Higher Education. Some of these include analyses of advertisements or computer games in media studies, films in historical studies, buildings in architecture, textbooks in education. Image acts as the catalyst or primary text for the written analysis. There are a range of
‘visual methodologies’ in data collection and data analysis, such as those used in visual anthropology, psychoanalysis, ethnographic approaches, content analysis, some of which include multimodal transcriptions. This kind of visual analysis is different to the written reflections accompanying predominantly visual texts discussed earlier. Often the image is not included in the text, such as in the analysis of a film or ethnographic data. Thus, the importance of description becomes paramount in order to evoke the visual text and the written piece needs to be independent to a certain extent of an accompanying visual. I will discuss an example of this type of text in a Humanities foundation level course and pose a question about how writing about the visual could possibly change the nature of the writing as well as standard written genres in Higher Education.

The students were required to attend an art exhibition at the National Gallery and to compare two paintings depicting Table Mountain in Cape Town, South Africa. What interests me about this assignment is the way in which writing about the visual seemed to ‘free up’ the genre of the academic essay to some extent. In many cases, there was a higher degree of authorial engagement than usual, more first person and experiential writing. Also, there seemed to be more ‘interference’ by the secondary school genres of ‘creative writing’ or ‘free narrative’. The following example is an extract from one student essay – the introduction and the beginning of the conclusion. In this extract, the writer represents himself as having different relationships with the visual and verbal modes. He takes the opportunity to show his proficiency in the written mode, whilst disclaiming his ability in the visual.
“Holiday time in …….. the Malay Quarter???”

Some person, who, personally, I don't believe had an assignment due, once said, “A picture is worth a thousand words” – so, technically! I should be 500 better off. That was not the way I was feeling upon, an hour after I had just awoken, stumbling in front of what would be the first of my two paintings. Namely James Ford’s ‘Holiday Time in Cape Town in the 20th Century’. I then proceeded to another aesthetically pleasing object and, peering over her shoulder, realised I was looking at what would be the second of my two paintings, Leng Dixon’s ‘Malay Quarter’, grinningly appreciating the ironic differences I had just observed between the two paintings, which will become apparent upon my conclusion. I will first give my immediate reaction and a basic description of the works, followed by highlighting my main theme the ‘real vs. unreal’, in particular with regards to the architectural and cultural representations, followed by the way and the reasons, I believe, Table Mountain is represented in each.

...

Well, when I initially started this assignment, and first viewed the first painting by Ford, the only comment my art-uneducated brain could come up with was “Wow, bet he’d be good at pictionary”, but my naivety soon turned into child-like fascination and I found myself unconsciously picking up the habits of seasoned head-tilters, examining the work from all possible angles and distances.

The student uses the first person rather than the distanced authorial persona of academic discourse. The writing is more experiential than theoretical. He uses idiosyncratic punctuation (the exclamation mark in the middle of a sentence, the three question marks at the end of the title) which resembles the intonation of colloquial speech, and perhaps also points to the influence of e-language. He also uses sentence fragments, such as “Namely James
Ford’s ‘Holiday Time in Cape Town in the 20th Century’, probably for emphasis. There is a predominance of adjectives, such as “grinningly”, which are not often found in academic discourse and provide a more personal and ‘colourful’ flavour to the writing. In the first paragraph, which is an introduction to the essay, he hints at the conclusion, without giving it away: “grinningly appreciating the ironic differences I had just observed between the two paintings, which will become apparent upon my conclusion”. This hinting at the conclusion fits in more with a storytelling genre, than an academic essay. He does, however, provide a coherent introduction which outlines a clear structure for the essay. A lot of the above essay is tongue-in-cheek, for instance the reference to art appreciators as “seasoned head-tilters”, making one realize how humourless a lot of academic discourse often is. Bernstein (1981) talks of ‘recontextualization’ of a discourse into a pedagogical setting. In this case, journalesque or ‘free narrative’ have not been recontextualized into the more authorially distant, more tentative and less humorous academic discourse described earlier. We get the sense that the student is not unfamiliar with the academic conventions, but is choosing to play with them to an extent in terms of language, register and content. Perhaps it is something about the analysis of the images in this particular task that enable him to do so.

Implications for writing pedagogy in Higher Education

I have argued that although the verbal is often prioritised over the visual in writing pedagogy in Higher Education, we need to take cognisance of the
different modal realizations of academic discourse. Exploring the affordances of modes and modal specialization with students seems to be a vital part of developing academic literacies practices. In Author (2006), I argue that certain functions like affect are developed in mode-specific ways, and can straddle both the visual and the verbal modes. Also, multimodal texts raise the important issue of assessment. New criteria are required for multimodal texts, such as appropriacy of modal usage in a particular context and for a particular purpose, linkages across modes (such as the visual-verbal linkages discussed above), innovation in the use of representational resources (Stein and Newfield 2006: 16).

Written reflections on visual productions or written analyses of visual texts can often involve a complex metalanguage. We need to redefine writing pedagogy through the development of metalanguages that will facilitate awareness and analysis of multimodal textual construction. Metalanguages across modes entail systematic technical knowledge of the ways semiotic resources are deployed in meaning-making. Familiarity with a basic ‘language’ for talking about images could make ‘reflections’ on visual compositions easier, and could facilitate more systematic analysis of multimodal texts. This paper advocates a pedagogy which opens up access to texts through making explicit how texts work, in the hopes that this will enable access to social and material resources (Cope and Kalantzis 1993: 63).

It is clear that using one mode to reflect on another can open up interesting spaces for reflection, as in the essay analyzing the images of Table Mountain
where writing about the images seemed to allow play with the academic writing conventions and the genre of the academic essay. Of course, as previously noted, this ‘freeing up’ of the student’s approach to the academic essay may also be linked to the experiential nature of the task or the particulars of the assignment. Although it is more usual in Higher Education for the written mode to be used to reflect on the visual, it would be generative to use different modes to reflect on each other. For instance, producing a satirical cartoon to comment on a piece of historical writing, or drawing an image to illuminate an aspect of a poem could be a fascinating exercise in the classroom.

The choice of how to represent data or create an argument presents complex choices about conjunctions of meaning and form in visual design. When creating texts, people bring together and connect the available form that is most apt to express their meaning at a given time. The writer selects particular design features to be significant at a particular point, in a particular context, creating a motivated conjunction of meaning and form (Kress and Van Leeuwen 2006). However, from an academic literacies perspective, it is important to realize that both the regularities of modes and the interests of people are socially shaped to realize convention and this takes on a particular form in the academic environment.

References


