Methodological Processes for Examining Melody in Sound

Betty J Noad  University of New England                         AARE NOA07339

1  Introduction

Sound is increasingly significant for making meaning (semiosis) in digital multimodal formats such as on the soundtrack of film, DVD and television. Contemporary digital texts such as video, film and television programs privilege sound features (voiceover, music, sound effects) to ‘design’ meanings and to ‘position’ listeners towards the interests of the authors of these multimodal texts. Indeed digital texts that persuade, such as television advertisements and film trailers, particularly feature sound to build a convincing message about a product, for consumers. Sound assumes a significant place alongside language and visual images in the digital texts of our multimodal landscape (Baldry & Thibault, 2006; Jewitt & Kress, 2003), and will be a crucial part of future texts that students must learn to comprehend and construct, and critically understand.

Syllabus documents in Australian states require students to learn about how multimodal texts are constructed in response to local and global communicative contexts, and to critically interpret and compose multimodal texts using the resources of language, image and sound. Multimodal literacy (Jewitt & Kress, 2003) requires students to learn about the meaning-making (semiotic) resources of multiple modes, enabling students to participate in the interpretation and design of future texts. The New South Wales 7-10 English Syllabus (Board of Studies NSW, 2001) and the Queensland Y1-10 English Syllabus (Queensland Studies Authority, 2005) demonstrate a significant response to the call for teaching multimodal literacy, requiring students to learn about the forms, features and structures of multimedia texts, and to learn to compose complex multimodal texts for a range of purposes. The Victorian Essential Learning Standards (VCAA & Authority, 2006) and the South Australian Curriculum, Standards and Accountability Framework (South Australian Department of Education, 2001) require primary and secondary students to learn to interpret, critically respond to and create multimodal texts.
To implement syllabus outcomes, teachers must access professional learning for effective teaching of multimodal literacy. For systematic and explicit teaching about the texts and technologies of today, teachers need comprehensive practical frameworks and metalanguages to plan their pedagogical practices (Unsworth, 2001). Currently teachers have access to comprehensive theories, frameworks and metalanguages to support teaching about linguistic features of texts (Halliday, 1985; Halliday & Matthiessen, 2004). Kress and van Leeuwen (Kress & van Leeuwen, 1996, 2006) have identified a framework and grammar for investigating the design of visual images. While syllabuses have made scant reference to the tools provided by Kress and van Leeuwen, teachers are gradually taking up their framework and metalanguage for explicit teaching about visual images, enabling them to become teachers of visual literacy (Noad, 2004).

Despite syllabus demands, teachers do not have access to frameworks or metalanguages for teaching explicitly about how meanings are made using the modal resources of sound. A review of current research literature highlights the dilemma: there is a paucity of research providing a cohesive theoretical framework for describing and explaining the textual resources of sound, or the role of sound in multimodal texts, for educational purposes. Indeed, any metalanguage for talking and teaching about the resources of sound is presented as tentative (McDonald, 2003; van Leeuwen, 1999) and tools to analyse the role of sound for making meaning together with other modes are only recently being developed (Baldry & Thibault, 2006; Jewitt & Kress, 2003; van Leeuwen, 1999).

Nevertheless, fragments of models, frameworks and vocabularies are emerging from research relating to multimodality and the area of sound, which have implications for teaching multimodal literacy. This paper outlines one research perspective on the theoretical modelling of sound as a social semiotic, and the semiotic resources of sound as conceptualised by van Leeuwen in his book, *Speech, Music, Sound* (1999), as it informs the methodological processes that are critiqued herewith. Following van Leeuwen, a conceptual and technical
The analysis of one semiotic resource, that of melody, is presented in this paper. *The Queen* film trailer is used as a model for analysis, to reveal how the resource of melody realises the expression of emotions and feelings. This analytic methodology is critiqued with reference to its technical demands for research, and its suitability for revealing sound semiosis in film.

2 The theoretical modelling of sound by van Leeuwen (1999)

Multimodal texts make meanings by integrating resources from multiple modes of communication. Each mode, such as language, or image, or sound, has ‘a regularised organised set of resources for meaning-making’. Following Halliday’s theorising of language as a social semiotic (1978) and his study of systemic functional linguistics (1985; Halliday & Matthiessen, 2004), other researchers have modelled the semiotic resources of image (Kress & van Leeuwen, 2006; O’Toole, 1994), and gesture (Martinec, 1998). In *Speech, Music, Sound* (1999) van Leeuwen theorises the mode of sound also as a social semiotic. van Leeuwen argues that the semiotic resources of sound are sufficiently developed to allow description of the way they systematically function to make meaning. In modelling sound this way, van Leeuwen is positioning sound in the developing discourse of multimodal semiotics, or systemic functional semiotics (Macken-Horarik, 2006).

In line with research in film, van Leeuwen conceptualises sound to include speech, music and other sounds, where ‘other sounds’ refer to sound effects not produced by humans or musical instruments (Bordwell & Thompson, 2001; van Leeuwen, 1999). van Leeuwen identifies six major domains of sound: ‘sound perspective, sound time and rhythm, the interaction of ‘voices’ (human or musical, individual or group), melody, voice quality and timbre, and modality’ (van Leeuwen, 1999 p. 9). Each of these aspects of sound is common to speech, music and sound effects. According to van Leeuwen, sound perspective, sound time and rhythm, melody and voice quality and timbre derive their semiotic affordances from their ‘materiality’ (1999 p. 125), that is from the concrete material resources (human, musical and non-musical instruments) used to produce these sounds. Van Leeuwen argues that the
semiotic resource of loudness realises the system of perspective in sound; the semiotic resource of rhythm realises the system of timing in sound; the semiotic resource of pitch realises the system of melody, and sound qualities are realised by all these semiotic resources as well as resources such as breathiness, or tension in sounds.

In providing a descriptive account of the semiotic resources of sound, van Leeuwen investigates how these resources can be used to ‘say and do things with sound’ and how to interpret their meanings (1999 p. 9). van Leeuwen theorises sounds as the actions and interactions of people, places and things in our environment, and that sounds are not things, nor can they represent things. Sound can present the hum of a car engine for example, but it cannot be the engine itself. As well as presenting action and interaction, van Leeuwen theorises that sound, like all other semiotic modes, can ‘create relations’ between the listener and the sound itself (1999 p. 12). For example, relations of ‘formality’ are created by the system of social distance, which is realised by the semiotic resource of voice quality (van Leeuwen, 1999 p. 24). Timing in music and speech can reflect the ‘affective relationship’ that people have with clocktime or the time of the metronome, whether they embrace or struggle with it for example (van Leeuwen, 1999 p. 50). Interacting sounds (of speech and music) reflect the degree of interactivity and involvement that people adopt in an interaction, and how that involvement is affected by all kinds of ‘unequal power’ relations (van Leeuwen, 1999 p. 71).

The semiotic resource of pitch, according to van Leeuwen, can present action and interaction in two ways. He argues that in speech and music the direction of the pitch movement, that is rising or falling pitch, indicates that communicative exchanges may continue (rising pitch invites continuity) or conclude (falling pitch indicates finality). Van Leeuwen suggests that this intonation finality and continuity forms a textural aspect of an ongoing sound event, by connecting and disconnecting parts within an audio text (1999 p. 101). However he suggests also that intonation finality and continuity can express interactive attitudes such as assertiveness or tentativeness, which are interpersonal meanings. As well, melody has the capacity to invoke
emotions and feelings, thus creating affective relations with the listener (van Leeuwen, 1999 p. 97). If, as van Leeuwen theorises ‘sounds are actions’ (1999 p. 93), or that sound presents and represents action, then it can be assumed that sound expresses emotions and feelings that relate to the action, in the same way that linguistically, verbs have evaluative and emotive features that express Affect (Eggins, 1994 p. 125)). Pitch is clearly a semiotic resource of sound that can realise *interpersonal* actions and interactions.

Van Leeuwen also draws on the work of Martinec (1996 in van Leeuwen 1999 p. 35) to use the term ‘express’ synonymously with the notion of ‘presenting’ and ‘representing’ significations with sound, according to whether the source of the sound is identifiable or not (1999 p. 35). Sounds present action and interaction if the source is visibly presented, whereas sound represents action and interaction if the source is invisible but can be identified. The sound of a door slamming can be presented with an image of a door slamming shut, and the same sound may later represent a door slam despite the source being invisible.

Following a systemic functional approach (Halliday, 1978. 2004), van Leeuwen develops a system network for each of the semiotic resources of sound. Each system network describes choices offered by a semiotic resource, that have semiotic value, or potential for making meaning. The next section of this paper elaborates on semiotic choices from Van Leeuwen’s ‘system network of melody’ (1999 p. 119), and models how technical analysis of pitch provides a descriptive account of these choices, and their potential meanings.

To explain how each of the semiotic resources of sound realise meanings, van Leeuwen extends his theoretical modelling to include an interpretive framework which engages the concepts of ‘provenance’ and ‘experiential meaning potential’ (1999 p. 46) and the concept of social context. The provenance of a sound refers to ‘associations or connotations’ invoked about the people and place where the sound comes from (1999 p. 210) and the experiential meaning potential of a sound refers to ‘our experience of what we
physically have to do’ to produce a particular sound (1999 p. 205). The experience and associations that we have with a sound generate meaning potential, but the actual meaning is then formed or clarified with reference to the context in which the sound is situated.

3 The Semiotic Resource of Melody

In his book *Speech, Music, Sound* (1999), van Leeuwen conducts discussion about the expression of emotions and feelings largely in relation to ‘melody configuration’ (van Leeuwen, 1999 p. 98), or the way that pitch is configured as melody in a sound act. The Harvard Dictionary of Music (Apel, 1971 p. 678) defines pitch as the ‘location of a musical sound in the tonal scale, proceeding from low to high’, that pitch is determined by the frequency (the number of vibrations per second) of the sound, and that while pitch as a physiological sensation depends to a small degree on other factors such as volume, these factors are ‘negligible from a musical point of view’. High pitch and low pitch are discernable in speech and music, and as van Leeuwen suggests, in sound effects (1999 p. 105). Variations in pitch configure the melody, or the ‘tune’ of the sound. To put it another way, the system of melody inscribes patterns of pitch onto rhythmic ‘phrases’ (structural units of sound), making them ‘melodic phrases” (van Leeuwen, 1999 p. 208).

This section of the paper emphasises two areas for demonstration and critique

- 3.1 and 3.2 examines the application of van Leeuwen’s methodological procedures for the segmentation of a sound text, rendering it for further transcription and annotation of pitch
- 3.3 examines the application of van Leeuwen’s concepts and tools for analysis of emotions and feelings in sound, as realised by pitch movement, pitch range, pitch level and melodic articulation.
van Leeuwen’s ‘System network for describing melodies’ (1999 p. 119) is included here to indicate interpersonal meaning potentials afforded by pitch, which are analysed in this paper:

![Diagram of van Leeuwen's 'System network for describing melodies'](image)

**Figure 5.7** A system network for describing melodies

### 3.1 The melodic phrase as a structural unit of sound

To describe how melody presents and represents emotions and feelings, van Leeuwen engages the concept of ‘melodic patterns’ (1999 p. 101), and the way that melodic patterns occur in a ‘melodic phrase’ (1999 p. 208). van Leeuwen theorises that ‘rhythm segments the stream of sound’ into discrete units (1999 p. 93), and uses this concept as an analytic tool to differentiate scalar units in the flow of measured sound, that is, sound that is measured by clock time or metronomic time. In relation to this paper on melody the ‘phrase’ is a critical structural unit for analysing speech and music, so is described in the summary that follows:
a ‘measure’ is the smallest structural unit of sound. Each measure is marked off by rhythmic, explicit ‘pulse’ (a note, a sound, a syllable) which carries ‘key information’ (van Leeuwen, 1999 p. 39), and is perceived to be of equal duration. In music a measure is a unit of meaning equivalent to a ‘bar’ (Apel, 1971 p. 513).

a ‘phrase’ is the fundamental structural unit, of varying length. Measures are grouped together in phrases as they are meant to be understood as ‘belonging together in some way’ (van Leeuwen, 1999 p. 41) or defining structural completeness. In speech each phrase usually comprises up to seven measures, and in music phrases are often four, and most often eight bars (measures) long (van Leeuwen, 1999 p. 47). Every rhythmic phrase has a ‘main pulse’ (van Leeuwen, 1999 p. 207). The boundary between phrases is usually marked by some kind of ‘audible break or change’, which is why phrases can also be known as ‘breath groups’ (van Leeuwen, 1999 p. 41).

a ‘move’ in the flow of sound is an intermediate unit of meaning. One or a number of phrases will make up a ‘move’ which is defined by van Leeuwen as a ‘turn in a dialogue, or a distinct section in a monologue, for example a verse in poetry, a chorus in popular music, a paragraph or a point in speech’ (1999 p. 208).

The ‘sound event’, which is the largest structural unit of meaning. The ‘sound event’ is an instance of communication by sound, such as a radio program, a public announcement, a musical piece presented in a concert program (van Leeuwen, 1999 p. 213).

In other words, the phrases, which are delimited by rhythm, are critical structural units which make up the moves in the ongoing stream of sound, or the sound event. Following van Leeuwen, this paper models how a sound event such as The Queen filmtrailer (Frears, 2006) can be segmented into structural units such as moves, phrases and measures, to facilitate further analysis of pitch.
The Queen filmtrailer was produced by Mirimax Film Corporation and released in 2006 prior the release of the feature film in 2007. The feature film is about how the British Royal Family, and the Queen in particular, responded to the death of Diana, Princess of Wales. Excerpts are taken from the filmtrailer, which lasts 2 minutes, to exemplify analytical concepts discussed in this paper. The Queen filmtrailer has five main sections:

1. The televised announcement of the death of Princess Diana
2. The Queen announces that no statement will be made by the Royal Family about Diana’s death, as the Queen considered that Diana was not a member of the Royal Family following her divorce from Charles.
3. Blair warns the Queen that her silence could be damaging to the Monarchy
4. The Queen resists pressure from the public that she speak to them about the death of Diana, but eventually acknowledges the need for her to do so
5. The Queen reflects on how she may need to change her approach to her position as Monarch, in line with the needs of her people

In the filmtrailer, the main characters that speak are the Queen and the Prime Minister Tony Blair. Other characters have minor speaking roles, such as the Duke of Edinburgh, Charles, and the Queen’s Secretary. Two instrumental pieces of music are included in the filmtrailer, Tune 1 is constantly repeated in Section 2, and Tune 2 is constantly repeated over Sections 4 – 5.

In speech, the frames for phrases are notated as [ ], measures are marked off by a single /, the pulses are **bolded** and *italicised*, and the main pulse is **CAPITALISED**

From Section 3, the speech of a television newsreader (a move in van Leeuwen’s terms) is segmented into measures and phrases:
Questions are being asked about WHY the Queen has not addressed her subjects at this TIME of national grief.

From Section 2, Tune 1 is notated as bars (measures) and a phrase:

Issues arising from segmentation of an audio text

van Leeuwen’s detailed definition of scalar structural units enables the researcher to complete segmentation of music and speech into measures or bars, and phrases. If the researcher cannot transcribe music using conventional notation, then expert musicologists need to be engaged for transcription of the audio text into bars and phrases. This researcher, whilst not a linguist, has been able to segment speech on The Queen filmtrailer into measures and phrases, as well as moves. However, to ensure reliability and validity of analytic research, the inclusion of credentialled musicologists and linguists for segmentation decisions is necessary.

3.2 The melodic phrase as a unit of meaning in sound

Phrases are structural units, analytic categories. If anything is going to be communicated however, suggests van Leeuwen, something will ‘have to be fitted into the frame’ of a phrase which can outline a message and therefore add meaning to it. A melody can add meaning to a phrase (1999 p. 93), as can timing, or loudness, or voice quality. According to van Leeuwen all phrases carry a melodic contour, or pattern, which defines the melodic phrase as a unit of meaning in sound. Linguists and musicologists have noted that the melodic phrase has a discernable beginning, middle and ending (Asaf’ev, 1977; Crystal, 1969; Halliday, 1967; van Leeuwen, 1999). It is the melodic contour of the melodic phrase that constitutes either a ‘sound act’ or one of
the repeated motifs or melodic patterns as a 'sound setting' (1999 p. 208). A sound act is defined by van Leeuwen as an 'individual unit of meaning' (1999 p. 212), and in this paper the focus is on how melody is inscribed onto a phrase to constitute a unit of meaning. The sound act is related to human action or social activities (van Leeuwen, 1999 p. 112), as distinct from melodic phrases which are repeated endlessly and become settings for human action (van Leeuwen, 1999 p.212).

van Leeuwen provides visual representation of the melody configuration of joy in speech and music, as observable patterns of melody:

![Image of melody configuration]

‘Wide pitch range at high pitch level. The melody rises, then falls sharply, then stays level (or descends slightly). Lively tempo. Pitch glides.’ (van Leeuwen, 1999 p. 95)

Van Leeuwen rejects a claim by linguists Fonagy and Magdics that the meanings of melodies are based on the words they go with, arguing that melodies are not slaved to words but ‘form an independent meaning system’ (van Leeuwen, 1999 p. 97) - exemplified by the situation in which the same words are sung with different melodies, yet realise quite different meanings.

**Issues relating to the visual display of melody in speech and music**

Van Leeuwen’s visual display of melody in speech and music in Fig 5.1 (1999, p. 95) is ideal. The display of melody in conventional musical notation is unambiguous, as pitch is identified by notes on a stave. The display of melody
in speech, following Fig 5.1, is ideal because the visual ‘contour’ identifies pitch in various ways (elaborated in 3.3), such as pitch movement, and by superimposing pitch contour onto a musical stave, information about pitch level (high or low) and pitch range (interval between high and low pitch) is unambiguous. The research analyst faces two issues in replicating the melody of speech in this way. A credentialled, practised linguist can only provide valid data by listening to audio text then transcribing pitch contour. To be able to superimpose that pitch contour onto a musical stave, or to express pitch contour in musical terms, requires particular expertise. Pitch analysis software, available free on the web, can analyse and display melody in speech in this way, but that technological analysis is contingent upon the researcher accessing a single track of recorded speech, free of interference from music, for instance. As this paper is written, the author/researcher continues to investigate technologies and expertise to objectively display melody in speech. In 3.3 the author has attempted to identify pitch patterns in speech in line with van Leeuwen’s model, recognising that they can only be described as ‘subjective’ descriptions requiring validation by experts.

The focus of the next section is to define and analyse the semiotic resources of melody with reference to van Leeuwen’s system network, and to model the application of these analyses to The Queen filmtrailer. As the focus of this paper is on methodological processes that investigate how melody realises interpersonal meanings, discussion is limited to patterns of pitch movement, pitch range and pitch level, which together, van Leeuwen asserts, realise emotions and feelings.

3.3 Melodic patterns in the whole melodic phrase

Whole melodic phrases are ‘configurations of different features’ such as pitch movement, pitch range and pitch level, which may ‘combine in a number of different ways and each contribute elements of meaning’ to the sound act constituted by the melodic phrase as a whole (van Leeuwen, 1999 p. 101).

Pitch movement: activation, stasis and deactivation
van Leeuwen states that ‘pitch movement’ creates the ‘melody’ or the ‘pitch contour’ of sounds (1999 p. 209), that is the pattern of high and low pitch inscribed onto the melodic phrase. van Leeuwen defines the main melodies: those which have a rising pitch and go ‘up’ are defined as ‘ascending’ (1999 p. 203); those which have a falling pitch and go ‘down’ are defined as ‘descending’ (1999 p. 204).

Van Leeuwen suggests that the meaning potential of pitch movement is experiential, relating to what we do when we produce it with our voice (1999 p. 103), identifying this link as ‘experiential meaning potential’. In describing the semiotic value of melodic patterns inscribed by pitch movement, van Leeuwen links experiential meaning potential to the interactive attitudes that pitch movement suggests. van Leeuwen concurs with Cooke (1959) for example, in arguing that ascending melodies are more active, energetic outgoing and dynamic than descending melodies, linking these attitudes to a physiological concomitant of singing wherein ascending pitch requires an increase in vocal effort, but descending pitch allows the singer to decrease the effort, thus descending pitch can relax and soothe listeners and make them focus on their thoughts and feelings (1999 p. 103).

Van Leeuwen suggests that pitch movement upwards realises ‘activation’ (1999 p. 203), that is, the more the pitch rises the more active and interactive the participants involved will be. The more the pitch falls, the more the participants will be deactivated and brought into some state of non-activity, which van Leeuwen defines as ‘deactivation’ (1999 p. 204). When there is no rise or fall in pitch and the pitch is level, such as a monotone sound, van Leeuwen defines this melodic pattern as ‘stasis’ (1999 p. 213). van Leeuwen (p. 104) provides Fig 5.5 and 5.6 as examples:
From Sections 4-5 of The Queen filmtrailer, the pitch in Tune 2 is ascending metaphorically activating communications as the Queen realises the need to communicate more with her people:

From Section 4, the Queen firmly states that she will not leave her grandsons to make a public statement about the death of Diana, deactivating communication with falling pitch movements:
But later in Section 5, the Queen considers that her position of silence may not have been an appropriate response to Diana’s death. Her voice rises in pitch, activating communication and thought on the possibility:

![What if my actions ARE damaging the Crown?](image)

**Issues relating to the analysis of pitch movement**

Identification and visual display of pitch movement in music is unequivocal. As discussed above, research analysts face the issue of objectively identifying pitch movement in speech to validate analytic decisions. Superimposing pitch movement onto musical stave identifies the extent of the ‘up and down’ movement, where the utterance starts from and where it is directed to. However as van Leeuwen’s descriptions require simply the identification of ascending or descending pitch, transcription on a musical stave may not be necessary.

**Pitch range: emotive expansion and emotive confinement**

van Leeuwen explains that whether ascending or descending, melodies can move in ‘small or large intervals’, such as large strides and energetic leaps, or restrained measured steps (1999 p. 105). Fig 5.5 demonstrates a melody moving up by large jumps, for example, whereas in Fig 5.6 the melody moves down by small steps (p. 12 in this paper); in other words Fig 5.5 reveals a greater range of pitch than Fig 5.6. ‘Pitch range’ is defined by van Leeuwen as the ‘scale running from maximally wide pitch range to a maximally narrow one (that is, to monotone)’ (1999p. 210). A ‘wide’ pitch range implies a wide interval, or difference, between highest and lowest pitch in a melodic phrase, a ‘narrow’ pitch range implies that the interval, or difference between highest and lowest pitch is narrow.
Van Leeuwen argues that the semiotic force of pitch range (or the experiential meaning potential) rests on what it is we do when we increase or decrease the pitch range. When we increase pitch range, he suggests, we are 'letting more energy out', and generalises that wide pitch range allows us to 'give vent to strong feelings'. In contrast, van Leeuwen argues that when we decrease pitch range we are 'holding more energy in', and that the narrow pitch range 'constrains the expression of strong feelings', whether as the result of a 'stiff upperlip' attitude, or because of modesty, or because we are paralysed with fear, or because we have no energy left for example(1999 p. 106). He cites Brazil et al in suggesting that wide pitch range conveys 'excitement', 'surprise' and 'anger' and that narrow pitch range conveys 'boredom' and 'misery' (1980 in van Leeuwen 1999, p. 106).

Indeed van Leeuwen argues that pitch range 'characterises the emotional temperature' of individual sound acts (1999 p. 106). An increased (wide) pitch range allows more room for the expression of feelings and attitudes, so is mostly responsible for realising 'emotive expansion', whereas a decreased (narrow) pitch range confines the expression of feelings and attitudes, so is mostly responsible for realising 'emotive confinement' (van Leeuwen, 1999 p. 205).

van Leeuwen provides the following examples (1999 p. 96) of narrow pitch range (Fig 5.4) and wide pitch range (Fig 5.3):

Please/ please/ don’t/ leave / me a/ LONE

You know my heart cries for you

Figure 5.4 The intonation of ‘anguish’ and its musical representation (in ‘My Heart Cries For You’, Ray Charles)
From Section 2 of *The Queen* film trailer, the Queen maintains that mourning should be conducted in private, and not subject to public scrutiny. From Section 5, she quietly considers her misjudgement. She speaks with a narrow pitch range, reflecting confinement of her feelings:

From Section 2, Blair expresses his exasperation with the Queen’s reasoning in relation to her silence, his voice rises in pitch and is expansive, reflecting the strength of his feelings:

From Section 2, Tune 1 has a narrow pitch range, reflecting the anguish of mourning, and dismay in relation to the Queen’s decision of silence:
Issues relating to the analysis of pitch range

Van Leeuwen’s conceptual analysis of pitch range to reveal the ‘emotional temperature’ of a text appears appropriate and logical, but technical analysis of pitch range is problematic for the researcher. Van Leeuwen does not define what constitutes ‘wide’ or ‘narrow’ pitch range in musical terms, and yet his visual display of pitch range indicates that it is a musical decision. While he exemplifies wide pitch range in Fig 5.3 and 5.5, and narrow pitch range in Fig 5.4 and 5.6, it is left to the researcher to define the range of pitch that can be described as ‘wide’ and vice versa, as well as identifying the critical musical point at which ‘wide’ pitch range becomes ‘narrow’ pitch range. Indeed van Leeuwen refers to ‘mid range’ pitch in an analysis of pitch range in *The Piano* (1999 p. 119), suggesting that a pitch range exists somewhere between ‘wide’ and ‘narrow’ that is significant enough to realise meaning for the listener, albeit that ‘mid range’ pitch is not featured on van Leeuwen’s system network. Precise definition of pitch range, as van Leeuwen or other researchers may conceptualise it, is required to enhance reliability and validity of analytic research decisions that describe melody in speech and music.

**Pitch level: dominance and danger**

Pitch level, like other aspects of pitch, relates to vocal effort: the higher the pitch, the greater the effort needed, and the more the voice is (literally and figuratively) ‘keyed up’; the lower the pitch, the less effort is needed, and the voice will sound ‘low key’ (van Leeuwen, 1999 p. 107). van Leewuen points out that speakers rarely use the complete pitch span their voices are capable of, and that people will adjust their pitch level for particular sound acts, for example, for expressing joy or excitement (1999 p. 107). Pitch level can also be related to a particular speech style or singing style, for example van
Leeuwen found that male newsreaders speak at a higher pitch level when they are on air than ordinary ‘low key’ speech (1999 p. 108), and Tagg noted that the typical voice of the male rock singer is characterised by high pitch (as well as loudness) as the singer will ‘raise his voice to an average pitch at least one octave above what he uses for normal speech’ (1990 p. 112 in van Leeuwen, 1999 p. 108). Examples of distinctive pitch level as conceptualised below by van Leeuwen, are provided from *The Queen* filmtrailer, where applicable.

Van Leeuwen distinguishes between high-pitched sounds that are loud or soft, and low-pitched sounds that are loud or soft, but notes that the meanings of these particular sounds are ‘strongly gendered’ (van Leeuwen, 1999 p. 111). In music or speech that is in a high key, or a high pitch register, ‘dominance’ is realised by a high voice which ‘leads’ by carrying the melody and is loud: in fact van Leeuwen defines ‘dominance as a ‘semiotic potential of a high pitch register combined with formal distance’ which derives from the fact that these sounds seek to cover a large territory (they are high and loud) (1999 p. 205). In opera, a soprano may dominate by singing loudly in her high-pitched register, but also a tenor may dominate by singing loudly in his (relatively) high-pitched register.

In *The Queen* filmtrailer, nearly all speech is amplified. Regardless of gender, two examples of high-pitched and loud speech - that dominates - can be identified. From Section 3, the Duke of Edinburgh is affronted when Blair suggests openly to the Queen that her actions are damaging the Monarchy. His voice is high-pitched and loud, demonstrating his belief that the Royal Family dominates:
From Section 2, the Queen declares that Diana is no longer considered a member of the Royal Family, following her divorce from Charles. Her voice is high-pitched and loud, intending her view to dominate those of others;

Speech and music in low pitch registers can also realise ‘dominance’ if the low pitch register is ‘combined with loudness’ (van Leeuwen, 1999 p. 205). A bass opera singer who sings loudly and at low pitch can dominate in sections of opera. An alto, or mezzo soprano (who sing at pitch lower than soprano) may not however, dominate sections of opera if they sing simultaneously with sopranos at the same levels of loudness. This distinction which implies that speakers and singers dominate the soundscape if they produce loud and low-pitched sounds, regardless of gender, is contestable.

As most speech in The Queen filmtrailer is high-pitched and loud, only one example of low-pitched, loud speech is identifiable. From Section 2, the Queen speaks loudly and at low pitch, commanding that no one in the Royal Family will speak to the public about the death of Diana:

When low-pitched sounds, which our experience tells us tends to be produced by large people or things, are made soft, they are described as realising ‘danger’ (van Leeuwen, 1999 p. 204). van Leeuwen defines ‘danger’ as a ‘gender stereotype’ in which ‘low pitch register combined with intimate distance or personal distance (softness) conveys a ‘dark’ and ‘dangerous’ femininity (van Leeuwen, 1999). Van Leeuwen also presents an example from opera to illustrate pitch level: in operas, he suggests, the tenor is often the...
hero and the bass is the villain (1999 p. 108). The implication here is that the bass with a low-pitched voice will be ‘dangerous’. In van Leeuwen’s system network of melody it will need to be soft, but in reality in opera the bass can be low and sing very loudly, yet remain villainous.

In The Queen filmtrailer, two examples of low-pitched, soft speech are identifiable, that realise danger, regardless of gender. In Section 2, the Queen speaks to Blair in a low-pitched soft voice, warning him of the danger in challenging her traditional view of mourning in silence:

In Section 4, Blair responds negatively to the suggestion by the Queen’s Secretary that the Monarch’s behaviour is willed by God. His voice is low-pitched and soft, but sounds a warning that the church has nothing to do with how the Queen should respond to her people:

Issues relating to the analysis of pitch level

Softness and loudness cannot be indicated on paper, so in this section the reader is expected to take on trust the examples of van Leeuwen (1999) and The Queen. van Leeuwen points out that meanings realised by pitch level and loudness are highly gendered, but The Queen examples contradict his assertions about meaning potentials, as distinguished in his system network. Research that further investigates the impact of gender on sound semiosis, in relation to pitch level and loudness, is required. At this point the research
analyst might not be able to validate these kinds of decisions by referring only to van Leeuwen’s work.

4 Conclusion

van Leeuwen’s theoretical modelling of sound as a social semiotic articulates with the current theorising of language and image, and is therefore helpful for thinking and talking about multimodal semiotics such as sound. The identification of key semiotic resources of sound also has educational implications, as it offers to schools a way of investigating and describing the behaviours of resources such as melody, and how they make meanings.

This paper examined only the semiotic resource of melody, and limited its investigation to how melody realises interpersonal meanings, as theorised by van Leeuwen. Under examination were the concepts and tools engaged by van Leeuwen to describe the resources of melody: how the audio text can be segmented for further technical analysis, and how a technical analysis of pitch features reveals the construction of emotions and feelings. Segmenting the audio text as a preparatory step in analysis of speech and music, in reference to van Leeuwen’s theorising, has been shown to be workable. The concepts that van Leeuwen engages to explain how pitch realises emotions and feelings, are reasonable and logical if considered in relation to his interpretive framework. The need to include expert musicologists and linguists in such analytic research has been pointed out in this paper, to ensure the reliability and validity of research decisions. Not all analytical tools that van Leeuwen engages for technical description of pitch are definitive, which makes further application of parts of his methodology problematic for the researcher. Further research and definition of pitch range, for example, is required before analysts can validate their research decisions. This paper flags the need for more research which informs the way we theorise the area of sound, and research which tests out the methods we might use to analyse the resources of sound. If teaching multimodal literacy is a social and educational imperative, then such exploratory research needs to progress.
References

McDonald, E. (2003). Sound as embodied behaviour: towards a model of music as a social semiotic system.


