'It's not what you said, it's how you said it!': Prosody and impoliteness

Jonathan Culpeper

1. Introduction

It is not an unusual occurrence that people take offence at *how* someone says something rather than at *what* was said. Consider this exchange between two pre-teenage sisters:

[1]   A: Do you know anything about yo-yos?
      B: That's mean.

On the face of it, Speaker A's utterance is an innocent enquiry about Speaker B's state of knowledge. However, the prosody triggered a different interpretation. Speaker A heavily stressed the beginning of "anything", and produced the remainder of the utterance with sharply falling intonation. This prosody is marked against the norm for yes-no questions, which usually have rising intonation (e.g. Quirk et al. 1985: 807). It signals to B that A's question is not straightforward or innocent. It triggers the recovery of implicatures that Speaker A is not asking a question but expressing both a belief that Speaker B knows nothing about yo-yos and an attitude towards that belief, namely, incredulity that this is the case − something which itself implies that Speaker B is deficient in some way. Without the prosody, there is no clear evidence of the interpersonal orientation of Speaker A, whether positive, negative or somewhere in between.

This example illustrates the fact that prosodic features play an important role in disambiguating messages, as research in communication has repeatedly demonstrated (e.g. Archer and Ackert 1977; DePaulo and Friedman 1998). Indeed, prosodic features do more than disambiguate messages: they can over-rule conventional meanings associated with linguistic forms, as is the case with ironic or sarcastic utterances. It is interesting to note the functioning of interpersonal communication in the written channel, i.e. without prosody. Kruger et al. (2005) investigated the constraints of the written channel in email communication, particularly regarding the communication of humour and sarcasm. In one study, they found that 75% of subjects who listened to statements successfully interpreted them, against 50% − that is, no better than chance − who read them on email (2005: 928). The relative paucity of communicative resources by which to secure an understanding, coupled with an egocentric tendency to "overestimate the obviousness of the fact that they are 'just kidding' when they poke fun or criticize" may lead people to "unwittingly offend" (2005: 934), thus contributing to phenomena such as 'flaming' (hostile, insulting computer mediated interaction).

Yet despite the importance of prosody in communication, the vast bulk of research on politeness or impoliteness pays woefully little attention to the role of prosody. The single exception of note is the work of Arndt and Janney (e.g. 1985, 1987), whose notion of politeness involves emotional support conveyed multimodally through verbal, vocal and kinesic cues. As far as studies focusing specifically on impoliteness or interpersonal conflict are concerned, prosody seems to attract at best a cursory mention (though see Culpeper, Bousfield and Wichmann 2003[1]; Culpeper 2005). If we turn to the literature produced by scholars of prosody, concerted attention is rarely given to im/politeness, largely because of the dominance, until recently, of the structuralist paradigm and the consequent focus on intonational phonology. An exception of note is research on acoustic features and politeness in Japanese (see, for example, Ofuka et al. 2000). To be fair, for several decades there has been a small but steady stream of studies on the attitudinal or emotive meanings conveyed by prosody (see, for example, the excellent summary in Murray and Arnott 1993). Work on attitudinal meanings clearly are of some relevance to im/politeness, not least because impoliteness involves an attitude (see below). However, much of this

work focuses on isolated utterances in lab-based studies. Only relatively recently have studies on prosody begun to take the role of in-context language use seriously (see, for example, the special issue of the *Journal of Pragmatics* in 2006 on the prosody-pragmatics interface).

There is no agreed definition of linguistic impoliteness (Locher and Bousfield 2008: 3). Even the term that is used for the notion is controversial (why not use 'rudeness' instead of 'impoliteness'?) (see the disagreements between Culpeper 2007 and Terkourafi 2007). My own definitions have evolved over the last dozen years, the last being:

> Impoliteness comes about when: (1) the speaker communicates face-attack intentionally, or (2) the hearer perceives and/or constructs behaviour as intentionally face-attacking, or a combination of (1) and (2).
> (Culpeper 2005: 38)

Whilst this has the merit of emphasizing that impoliteness arises in social interaction (it is not simply something the speaker does), it nevertheless tacks the notion of impoliteness on to the notion of 'face-attack'. But that simply transfers the explanatory load on to another notion that is itself controversial (see, for example, Bargiela-Chiappini 2003) and may well not cover all cases of impoliteness (see Spencer-Oatey 2002). Below I give my current definition:

> Impoliteness is a negative attitude towards specific behaviours occurring in specific contexts. It is sustained by expectations, desires and /or beliefs about social organisation, including, in particular, how one person's or a group's identities are mediated by others in interaction. Situated behaviours are viewed negatively − considered 'impolite' − when they conflict with how one expects them to be, how one wants them to be and/or how one thinks they ought to be. Such behaviours always have or are presumed to have emotional consequences for at least one participant, that is, they cause or are presumed to cause offence. Various factors can exacerbate how offensive an impolite behaviour is taken to be, including for example whether one understands a behaviour to be strongly intentional or not.

This chapter both argues that prosody plays a key role in triggering evaluations that an utterance is impolite, and, moreover, aims to show some of the ways in which this happens. Much of what I say will also have relevance for the study of politeness. I start by considering the scope of the term 'prosody', and by noting key issues and trends in the field of prosody. In writing this overview, I assume that the readers of this chapter will know relatively little about prosody. My next step is to review relevant work (largely undertaken by pragmatics scholars) on prosody and im/politeness. Following on from this, I examine metapragmatic comments concerning potentially impolite utterances (see, for example, the title of this study). The point here is to demonstrate that prosody plays an important role in the 'lay person's' understanding of impoliteness. Finally, I take the crucial next step of examining how prosody works in context to trigger evaluations of impoliteness in naturally occurring data. I will analyse extracts from the singing talent show *Pop Idol*.

2. About Prosody

How does 'prosody' differ from 'paralanguage', and what is it? Both terms refer to vocal effects that do not constitute the verbal aspect of a message – the words and the semantics and grammar that accompany them – but which accompany it (or, less often, stand alone as the message themselves, when people make vocal noises but do not articulate words). Prosody is often taken to refer to more local dynamic vocal effects, variations in loudness, pitch, tempo and so on. This usage is also consistent with the long-standing use of the term to describe the rhythm of verse. Paralanguage, a term attributed to Trager (1958), refers to more general vocal characteristics – voice setting, voice quality, characteristics such as whining, laughing, whispering, etc., as well as vocalizations such as "uh-huh" or "mhm". However, usage of these labels has not maintained such a distinction. Paralanguage, for example, is often taken to refer to anything 'beyond language', including kinesic aspects, such as gesture, gaze direction and posture. In this paper, I

will use prosody to refer to any vocal effect, leaving paralanguage as a superordinate term. Having said that, it is not a helpful starting point for an analysis to say that prosody is *any* vocal effect – what particular aspects of vocal effects will be attended to in the prosodic descriptions? Also, the local vs. general distinction touched on above is a useful one which we should not lose sight of. I will elaborate briefly on each of these in turn.

Prosody in this chapter refers in particular to:

Timing. Includes: speech rate (e.g. segments, syllables or words spoken per second), duration (e.g. of syllables, tone units, utterances), rhythm (i.e. patterning of accented syllables); pauses (frequency and length).

Loudness. Measurable in Decibels.

Pitch. The frequency of the vibrating vocal folds − the "fundamental frequency" (F0) (measurable in Hertz). Includes: pitch range (i.e. the range of pitch values produced by a speaker during a tone unit, utterance, speech event, etc.), pitch contour direction (e.g. fall, rise, fall-rise, rise-fall, level), pitch contour gradient (i.e. the steepness of the contour (the rate of change in pitch per second)).

Nucleus. The most important accented syllable(s) in the tone group, usually louder, longer and of higher pitch (and the last accent if the tone group contains more than one). Includes: nucleus prominence (i.e. degree to which the nucleus is made prominent through, primarily, loudness and pitch height), nucleus placement (i.e. the positioning of the nucleus in the tone group).

Voice quality. The harmonic overtones or auditory colouring that accompanies the fundamental frequency (produced by the settings of the larynx and supralarynx and overall muscular tension). It includes voice qualities such as: breathy, whispery, creaky, harsh, falsetto.

It needs to be stressed that the fact that I have presented five *separate* groups is a matter of presentational and analytical convenience. Prosody is a composite of acoustic features. Indeed, the five groups are not only interrelated but also sometimes themselves composites of some of the other groups. For example, a voice quality such as breathy involves low pitch; a tone group nucleus is primarily comprised of loudness and pitch. Also, note that the notion of nucleus assumes another concept, namely, the tone group. There is no hard and fast way of defining a tone group. It usually contains one nucleus, forms an unbroken rhythmical sequence and is bounded by pauses. Wichmann (2000: 4) states that it contains 4–6 syllables on average, and Quirk et al. (1985: 1602) state "around five to six words, or two seconds".

Descriptions of prosody can proceed along three lines: (1) auditory descriptions, i.e. using descriptive words and transcription conventions to represent what the analyst hears, (2) acoustic descriptions, i.e. using instruments to represent the physical properties of the sound, and (3) articulatory descriptions, i.e. using descriptive words to represent the physical mechanics involved in the speaker's production of the sound. This chapter will not engage in articulatory descriptions. Acoustic descriptions, deploying the software program *Praat* (version 5.0.40; Boersma and Weenik 2008) (available at: http://www.fon.hum.uva.nl/praat/), will be used for the key segments of talk under scrutiny in the final sections of this chapter. Such descriptions have the merit of being relatively objective. They are also relatively accessible to readers. For example, seeing the visual representation of pitch movement as a line going up or down in visual space is more accessible than a transcription system using alphabetic letters (e.g. the autosegmental system). It is also more delicate, as the slightest variation in pitch will affect the line. There are, however, limitations. A practical limitation is that software programs can only handle relatively clean recordings. This rules out analysing a recording made in a busy street or on a crackling telephone. The data analysed in this chapter is produced in a studio, and thus the audio quality is relatively high. Another limitation is that not all vocal aspects are equally amenable to instrumental analysis. Until relatively recently, the analysis of voice quality has been limited to auditory descriptions (e.g. breathy, creaky), because of the particular technical difficulties involved (for recent advances involving electroglottography, see http://www.ims.uni-stuttgart.de/phonetik/EGG/frmst1.htm).

Furthermore, a more fundamental limitation is that there is no necessary correlation between what an instrumental analysis represents and what someone perceives. The analyst must decide on what is relevant. Also, an instrumental analysis does not necessarily display the context against which acoustic features are perceived. Instrumental analyses are usually focussed on relatively short segments of speech. For example, somebody's pitch range may seem unusually high, yet this could turn out to be a general feature of their voice. For these reasons, I will supplement my analyses with auditory descriptions.

The example given in the first paragraph of this chapter illustrates local and (general or global) issues. Intonational contours – falls, rises and so on – are, at least in part, conventionalised local events and are contrastive (e.g. a fall contrasts with a rise), and as such are relatively easy to identify and to interpret. Indeed, this is the traditional end of prosody that some treat as part of a grammatical system, studying, for example, how certain sentence types and intonation contours correlate, and how the intonation nucleus is associated with the information focus of the sentence. This area of prosody, focussing on the more general, might be described, following Roach (1991), as 'linguistic form-based prosody'. Example works include Halliday (1967) and Ladd (1996). Other vocal phenomena are more clearly gradient. One speaks fast or slowly, for example, to some degree. The key point here is that what counts as fast or slow, high pitch or low pitch, and so on is relative. Relative to what? It could be relative to the local context, for example, the rest of the speaker's utterance or the immediately preceding speaker's utterance. It could be relative to the general context, for example, what is usual for that type of speaker (e.g. a man or woman, young or old person). It could also be relative to an aspect of the context somewhere between global and local, such as what is usual for that speech activity or event (e.g. increased loudness addressing a public meeting). Two areas of prosody focus on the more local context. One, again following Roach (1991), is 'interpersonal prosody' (an area which encompasses affective or attitudinal prosody), and involves the communication of (emotional) attitudes concerning the speaker, the speaker's message and/or the speaker's addressee (e.g. the use of high pitch to achieve the function of hedging). One classic example is O'Connor and Arnold ([1961] 1973). Designed for learning of English, the idea is to tell students what attitudes are conveyed by a particular type of intonation contour in conjunction with a particular grammatical structure. Thus, for the intonation contour of the first example of this chapter, a rapid high rise followed by a fall and a tail in the context of a yes-no question, they suggest the attitude conveyed is "impressed, challenging, antagonistic". However, the attitudinal correlations of this work seem to be based on intuition, and it also assumes a rather limited view of communication, in which prosody and grammar are simply a code for attitudinal information. In contrast, the final area of prosody studies - which we might label 'interactional prosody' – focuses on context, particularly the local context. The landmark work here is Couper-Kuhlen and Selting (1996). Interactional prosody involves either or both (a) the management of turn-taking (e.g. turn-completion associated with a drop in pitch) (e.g. Zuraidah and Knowles 2006), or (b) contextualisation cues (e.g. Brazil 1985; Gumperz 1982) (the use of prosody to project the contextual implications of an utterance).

In fact, even relatively conventionalised features, such as intonation contours, are not immune to gradience and relativity – all prosodic cues are gradient and relative (cf. Crystal and Davy 1969: 108; Arndt and Janney 1987: 227). Again, recollect the example with which this chapter opened. The fact that the utterance was taken to be 'mean' by Speaker B was probably triggered by the relationship between the yes-no question and its prosody (a fall). However, it is also the case that the fall commenced from a particularly high point and was particularly rapid. This begs the question of 'high' and 'rapid' relative to what. It is precisely the gradience and relativity of prosody that makes it crucial to account for the pragmatic inferencing that underpins its role in communication. Relevance theory (Sperber and Wilson 1995), for example, can account for which aspect of context is most relevant: essentially, that aspect of context which, in conjunction with the prosody, can maximise cognitive rewards (e.g. new information) for the least (processing) effort.

3. Politeness and Prosody in Communication

As I have remarked in the past (1996: 355), impoliteness is, at least to some extent, parasitic on politeness. It thus makes sense to consider the literature on politeness and prosody. The most influential work on politeness, Brown and Levinson (1987) (hereafter B&L), makes a few remarks in this connection. Some are scattered in their linguistic output strategies:

Negative politeness:
Output strategy: Question, hedge
Perhaps most of the verbal hedges can be replaced by (or emphasised by) prosodic or kinesic means of indicating tentativeness or emphasis. … In Tzeltal, there is a highly conventionalised use of high pitch or falsetto, which marks polite or formal interchanges, operating as a kind of giant hedge on everything that is said. (Elsewhere (Brown and Levinson 1974) we have argued for the universal association between high pitch and tentativeness). (B&L 1987: 172)

Positive politeness
Output strategy: Exaggerate (interest, approval, sympathy with H)
This is often done with exaggerated intonation, stress, and other aspects of prosodics (B&L 1987: 104)
Output strategy: Presuppose/raise/assert common ground
The personal-centre switch [the speaker adopts the position of the hearer] … can be carried out in prosodics as well: both A's and B's utterances could be expressed with 'creaky voice' (very low pitch and a constricted glottis), where the prosodics of giving comfort is the same as (or a metaphor for) the prosodics of asking for sympathy (B&L 1987: 119)

Off-record
Output strategy: Be ironic
… there are clues that his intended meaning is being conveyed indirectly. Such clues may be prosodic (e.g. nasality) … and in English ironies they may be marked by … exaggerated stress. (1987: 221–222).

More substantially, towards the end of the book they comment on prosody and phonology and politeness, and make the following argument and prediction:

Now our point here is not simply that there are correlations of prosodic or phonological features with social contexts, but rather that there are rational reasons why these particular features are used in these particular circumstances. For instance creaky voice, having as a natural source low speech energy, can implicate calmness and assurance and thence comfort and commiseration, attitudes not suitably expressed in negative-politeness circumstances. On the other hand, high pitch has natural associations with the voice quality of children: for an adult thing to use such a feature to another adult may implicate self-humbling and thus deference (Brown and Levinson 1974). We predict therefore that sustained high pitch (maintained over a number of utterances) will be a feature of negative-politeness usage, and creaky voice a feature of positive-politeness usage, and that a reversal of these associations will not occur in any culture. (B&L 1987: 268)

It is not difficult to find examples of Brown and Levinson's (1987) strategies that deploy prosody in the way they suggest. But it is also fairly easy to find the same specific prosodic features doing completely different things. However, their claims are carefully worded – "can implicate", "may implicate" – to accommodate such exceptions. Regarding the claim about high pitch always being a feature of negative-politeness and creaky voice always being a feature of positive politeness, this is an empirical question which, as far as I know, has not been addressed. However, the "rational reasons" put forward are not in themselves convincing. High pitch has associations, and arguably "natural" ones, with anger (specifically, hot anger or rage) (cf. Murray and Arnott 1993: 1103–4, 1106). This is hardly the stuff of "self-humbling" and "deference". Also, glancing ahead to section 4, there is evidence that high pitch associated with children and in particular contexts is considered "whining", something which suggests not positive but negative implications. Regarding their claim of universality in the final sentence, the lack of cultural

variation with respect to a number of different prosodic features is in fact supported by Frick (1985: 414–5). However, more recent research has highlighted clear cultural differences in the politeness perception of prosody. For example, in his work on 'pressed voice' (giving rise to a rasping quality due to tensed vocal folds) Sadanobu (2004) shows that it is considered part of deferential politeness in Japanese. In contrast, Shochi, Auberge and Rilliard (2007) show that the same voice quality is viewed negatively by Americans and the French.

As I indicated in section 2, prosodic features are unlike traditional linguistic features in that they are all gradient, at least to some degree. The key question, then, for understanding the role of prosody is: what is the norm (or norms) against which prosodic features are marked? Let us turn to the only work that both deals with politeness *and* provides a thorough treatment of prosody, namely, Arndt and Janney (1987). The thrust of their approach to communication is very different from Brown and Levinson (1987). According to Arndt and Janney (1987: 248, et passim), "utterances become 'meaningful' – by which we mean interpretable – only through the interaction of verbal, prosodic, and kinesic actions in context". Consistent with this position, they point out that "[i]t seems that it is not the simple occurrence per se of prosodic effects that is significant in interpersonal communication, but the distribution and intensity of these effects in relation to the acoustic bass line (cf. Crystal and Davy 1969: 108)" (1987: 227–228). In the main, they are concerned with multi-modal emotive communication. They argue that people, as opposed to social situations and their norms of appropriacy, are the locus of politeness, and that we should "focus on cross-modal emotive behaviour as a means by which politeness is negotiated" (1987: 377). It is not pertinent to this chapter to evaluate their understanding of politeness, though I briefly note that whilst "appropriacy-based approaches to politeness" are considered "too vague" as an approach to politeness (1987: 376), they deploy the notion of "frame" (Goffman 1974), which includes situational factors, to explain how the "emotive interpretations of normal behavioural patterns" (1987: 337) can be modified. In other words, if politeness is achieved through emotive behaviour and if emotive interpretations are sensitive to situational norms, then politeness is sensitive to situational norms too. Wisely, Arndt and Janney (1987) generally refer to the notion of "supportiveness", rather than politeness. The notion of supportiveness follows Brown and Levinson's (1987) notions of positive and negative face wants:

> A supportive speaker smooths over uncomfortable situations, or keeps delicate situations from becoming interpersonally threatening, by acknowledging his partner's claim to a positive public self-image. … A supportive speaker tries to minimise territorial transgressions and maximise signs of interpersonal acceptance. (Arndt and Janney 1987: 379)

The important thing for Arndt and Janney is then to forge a link with emotion:

> The basic idea which was introduced by Berger and Calabrese (1975), and subsequently developed into an axiomatic system by Bradac, Bowers and Courtright (1979, 1980), is that the listeners' emotional security depends to a large extent on the assertiveness, value-ladenness and intensity of his [*sic*] partner's behaviour. … high levels of emotional security produce increases in liking and intimacy and decreases in reciprocity and information seeking; low levels of emotional security produce decreases in liking and intimacy and increases in reciprocity and information seeking (cf. Berger and Calabrese 1975: 103–107). Emotional security, which in part stems from the belief that one can predict or explain one's partner's behaviour (cf. Bradac, Bowers and Courtright 1980: 213), is highest when face-needs are met, and lowest when they are threatened. (Arndt and Janney 1987: 380)

The chief merit of Arndt and Janney (1987) is a systematic and detailed discussion of how words and structures, prosody and kinesic features interact and create meaning in communication. Of particular relevance to the concerns of this chapter is their discussion of attitudinally marked prosody. They argue that it is unexpected prosodies that trigger the search for attitudinal interpretations, including "interpersonal interpretations, e.g., the customer wants to insult me (he's being arrogant, impolite,

dominating, etc. on purpose, who does he think he is?)" (1987: 273). Similarly, the near final sentences of Ofuka et al.'s (2000: 215) paper on prosodic cues for politeness in Japanese are worth noting:

> People appear to be very sensitive to unnaturalness by their standards and this listener-specific sensitivity may bias politeness judgements. A single extreme value for any acoustic feature (e.g., very fast speech rate) may reduce perceived politeness, but this will differ listener by listener.

Attitudinally marked intonation contours Arndt and Janney define as those that are not clearly motivated by syntactic considerations (1987: 273–274). For example, a declarative has the expectation of a falling intonation contour, and so a rise would be attitudinally marked. They suggest the following set of possibilities:

(1)   *rising pitch* together with declarative, imperative or wh- interrogative utterance types would be considered attitudinally marked;
(2)   *falling pitch* together with all other interrogative utterance types would be considered attitudinally marked;
(3)   *falling-rising* pitch, as a mixed contour, would be considered attitudinally relevant, regardless of the utterance type with which it is combined;
(4)   *all remaining combinations* of pitch direction and utterance type – i.e., the so-called 'normal' ones, grammatically speaking, would be considered attitudinally relevant only in conjunction with other types of cues or cue combinations.

<div align="right">Arndt and Janney (1987: 275)</div>

The first three items all involve "contrastive patterning", which Arndt and Janney view as "central to emotive communication" (1987: 369). They also discuss "redundant patterning", but "redundant patterning amplifies verbal messages", whereas "contrastive patterning modulates or modifies them" (1987: 369). The use of contrast between what is said and how it is said in the case of irony or sarcasm is an illustration of the latter.

Arndt and Janney (1987) propose a neat scheme. However, what particularly concerns me is that prosodic features can be marked for various reasons, not just syntactic. One of the key aims of the up-coming analysis is to show how other aspects of context, and in particular the co-text but also the situation, contribute to establishing a speaker's prosody as marked.

4. The Importance of Prosody for Impoliteness: Meta-impoliteness Comments

Some relatively recent major works in the area of politeness have argued that traditional approaches to politeness have applied pseudo-scientific politeness theories and categories to certain social behaviours (e.g. Eelen 2001; Watts 2003). They argue that research should focus instead on the lay person's conception of politeness as revealed in participants' own discursively-constructed and dynamic interpretations of local and particular encounters, and especially in their use of and disputes about the meanings and applicability of terms such as *polite* and *politeness* for such encounters (the same argument is made of impoliteness/ rudeness in Watts 2008). Whilst there are difficulties in constructing a general approach to politeness based solely on such a focus, it is certainly the case that any approach to either politeness or impoliteness needs to take account of the lay person's conceptions, as revealed, for example, through their use of terms. Indeed, metapragmatic comments and metalinguistic labels are important sources of evidence about how people understand politeness and impoliteness. They are also an important means of validating theoretical categories and their application. In this section, I will examine 'everyday' meta-impoliteness comments in order to support the idea that prosody can play a crucial role in triggering the interpretation of impoliteness, and also to give some limited clues about the ways in which it can do that.

My meta-impoliteness comments are drawn from weblogs, where they are fairly frequent. All

weblogs, as far as I can tell, related to the North American context. A preliminary survey of the data suggests that utterances are sometimes perceived as impolite because of the particular 'tone of voice'. We need to be a little cautious here, as 'tone of voice' is not always used to refer to prosodic features; it is sometimes used to refer to register issues (for example, language being too colloquial). Nevertheless, it is often used of prosodic features, and this is illustrated in [2] (Note: I have made no attempt to correct the typos and other infelicities in the weblog data):

[2] Anyone have any ideas on how to quell the nasty tone of voice? This seems to a constant struggle in our house. Instead of using a polite tone of voice to ask a question or make a request, there is that demanding, whiny voice, especially when my kids talk to each other. I can remember my parents saying to me 'It's not WHAT you said, it's HOW you said it.' Boy, do I know what they were talking about.

http://midvalleymoms.com/index.php?q=tone_of_voice

The 'tone of voice' is negatively evaluated, being described at the outset as "nasty". Because it is contrasted with "a polite tone of voice", it is by implication an impolite tone of voice. Moreover, it is impolite in a specific pragmatic context, namely, that of asking a question or making a request. In this context it becomes "demanding", probably because it conflicts with the likely normal and/or prescribed social organisation of the family in this culture, namely, that only parents have the power to demand. Note that even when the "kids talk to each other", this tone of voice is proscribed. Regarding the specifics of the voice, we learn that it is "whiny", an auditory description that suggests the voice is markedly high-pitched and enduring.

As [3] illustrates, an impolite tone of voice is not restricted to children as perceived by parents:

[3] It'slike fingernails on a chalkboard. You know the tone of voice. The tone of voice that infers 'You're a Moron' without even saying it. It's the tone of voice that leaves children whimpering. When I hear it coming out of another parent's mouth, I cringe. I'm not judging them - I feel for them. …

It's almost impossible to avoid that tone of voice when we're exasperated. But that tone says so much more than the actual words we're screaming. That tone says 'Why am I wasting my breath on you.' It says 'How did I get stuck with such a moron?' It says 'I'm embarrassed to be associated with you.'
http://nwitimes.com/parent/blogs/fatherknowsnothing/?p=19

Here, a parent evaluates the tone of voice of other parents negatively. Whilst no metalinguistic impoliteness label is used, there seems little doubt that the implications of the tone of voice, as indicated in quotation marks, conflict with emotionally-sensitive norms and beliefs relating to the mediation of people's identities – they are face-attacking. The writer also alludes to a negative emotional reaction on the part of the target ("whimpering") and third-party observers ("cringe"), further reinforcing the idea that an impoliteness attitude has been evoked. There is little detailed prosodic information. Being like "fingernails on a chalk-board" suggests very high-pitch.

A particularly common way of picking out prosody as the main trigger for evoking an impoliteness attitude is to say, as illustrated in [2] above, "It's not WHAT you said, it's HOW you said it", or some similar variant. Consider the examples given in [4] and [5]:

[4] Oh and for that manager, my friend noticed that the area where she was sitting was wet- yes water under the table. She came to wipe the floor herself but how she said 'its okay now' was so patronizing and condescending that we should have taken it as a sign. We should have not stayed to order.
http://winecountry.citysearch.com/review/35188398

[5]   After explaining the part to your first customer representative regarding the time off I had to take she responded by saying 'what would you like me to do about that sir?'. This is a not what she said but how she said it moment. I didn't even get to get to the point where I explained how I had been handled over the phone via Advance Video because at this point I was being challenged by your own representative in a manner akin to vermin

<div align="right">http://www.usualdosage.com/?p=100</div>

Examples [4] and [5] are exactly like the example with which I opened this chapter. There is nothing in the words or structures "its okay now" and "what would you like me to do about that sir?" that could be taken as a means of evoking an impolite attitude (in fact, the second utterance concludes with a deferential marker); neither is it the case that the particular context of those words and structures makes the evocation of such an attitude likely (it is not difficult to imagine them as polite enquiries). It clearly must be the prosody that interacts with the words and structures and their contexts that is at issue (in the case of [4] it is possible that kinesic cues were involved). No specific information is given about the prosody. One might suspect there to have been some mismatch between the expected prosody of these utterances in such contexts and what is realised. For example, the grammar of both predicts a rising intonation contour, but perhaps a fall of some kind was given (as in example [1]); or maybe the pitch range, nucleus prominence, duration of syllables or some other aspect was odd. Regarding the specifics of the impoliteness attitude evoked, both examples are similar: "patronizing and condescending" and "in a manner akin to vermin" indicate that the prosody infringes equity rights concerning fair treatment by others (see Spencer-Oatey 2008).
The final example of this section, [6] below, is similar to the two examples above.

[6]   I could go on and on, but to get the full effect, you had to see not just *what* Mary said, but *how* she said it. She was dripping with contempt and sarcasm, parroting anything said by the other panelists in a teenage sing-song imitation complete with the liberal use of air quotes.
       http://www.huffingtonpost.com/arianna-huffington/russert-watch-the-mary-m_b_15990.html

The implication of "to get the full effect" is that only a partial effect is available without prosody: the prosody guarantees the interpretation. The writer is quite specific about the aspects of the impoliteness attitude evoked. "Contempt" is reminiscent of the patronising and condescending prosody of the previous two examples, but we also have "sarcasm" and "parroting". Researchers have stressed the importance of prosody in communicating sarcasm, and pointed to acoustic features such as a lowering of pitch (the mean fundamental frequency) (see Cheang and Pell 2008). The prosody of sarcasm and mimicry are discussed and illustrated in Culpeper (2005).
What these meta-impoliteness comments establish is that prosody plays a role in evoking impolite attitudes. The mere presence of a particular 'tone of voice' in a particular context can be enough to act as a cue. In the earlier examples discussed, very high-pitched voices were viewed negatively. But, of course, this is not the only kind of prosody that can evoke impoliteness. Also, a very high pitched voice is not in itself impolite – it could, for example, signal excitement. It is the use of a very high pitched voice in a particular context, such as a child requesting something from a parent, that evokes impoliteness. The examples relating to how something was said show the crucial role prosody has in evoking impoliteness despite the words not doing so. Given that examples [4] and [5] are public service encounters with a waiter or telephone customer representative and [6] is a talk show with Mary Matalin, a spokesperson for the ex-U.S. Vice-President Dick Cheney, it seems likely that the parties responsible for the behaviour are con-strained from going fully on record, i.e. evoking impoliteness through both what is said and the prosody. It is much easier for them to deny relative and gradient meanings conveyed by prosody than the formal meanings of 'what is said', even though the prosodically evoked meanings may be considered blatant by participants. The fact that they are blatant yet not part of the formal record means that they fit the strategy for evoking impoliteness labelled 'off record impoliteness' in Culpeper (2005).

5. Prosody and Impoliteness in Interaction: An example from Pop Idol

5.1 Pop Idol and Susie's performance

None of the publications discussed in the previous section, nor indeed most of those referred to in this chapter, actually analyze language as part of extended social interaction and part of a particular context. Instead, they prefer to make claims about decontextualized and relatively short utterances, drawing on intuitions, either those of the analyst or those elicited from informants in experimental conditions. This section aims to fill that gap. In particular, I aim to reveal how locally constructed norms impact on politeness and, especially, impoliteness. An important source of evidence for my interpretation of utterances will be understandings displayed by their targets. Unlike the ambiguous example with which this chapter opened, my focus will be on the role of prosody as one of a number of signals which combine to exacerbate further the offence of utterances (i.e. the fact that it is offensive to a degree is already clear without the prosody).

My data is taken from *Pop Idol – Raw Talent* , a selection of extracts from a British television series which debuted on ITV on October 5, 2001. It is a talent show, in which a panel of judges initially, and later also viewer voting, determines the best singer, or 'pop idol'. It has been very successful. A second series followed in 2003, and the format has been franchised around the world as the "Idol" series (e.g. *American Idol*). It is also very similar to the series *X Factor*. It is not, however, an ordinary talent contest but an exploitative one. Exploitative chat, quiz or talent shows are structured to maximise the potential for face loss (see Culpeper 2005). Contestants expose their ability or lack of ability in public, in front of both the immediate audience comprised of the judges, television crew, etc. and the distant TV audience. In the first round of auditions, they have to sing standing in the middle of a fairly large room with no backing music or microphone, and with the four judges scrutinizing them from behind a desk – a situation likely to bring out the worst of performances from all but the steely-nerved. The evaluation of the performance, and thus performer, is not simply a numerical score or a ranking process (for example, identifying the top three) but a verbal assessment. To carry off a critical assessment without damaging face would require consummate tact in this context. But, of course, from another perspective, it is an opportunity to evoke strong impoliteness. Simon Cowell, a judge on all of these programmes, has developed a reputation for his acerbic remarks on contestants' performances. His autobiography is called *I Don't Mean to be Rude, but: The Truth about Fame, Fortune and my Life in Music* (Simon Cowell 2004), and his "rude" remarks have been collected in *I Don't Mean to be Rude, but: Simon Cowell's Book of Nasty Comments* (Tony Cowell 2006).[2] In fact, the verbal treatment of contestants on *Pop Idol*, and in particular the remarks of the judges Simon Cowell and Pete Waterman, was the subject of critical comments made by the British Members of Parliament Austin Mitchell and Jim Sheridan. Having said all this, there are contextual mitigating factors for potential impoliteness on *Pop Idol*, namely, the fact that it is all a game and that the judges may well just be playing the role of being nasty. However, in Culpeper (2005) where I analysed a quiz show, I argued that there is evidence that contestants took offence nevertheless, and such evidence is even clearer in *Pop Idol*, whose contestants are often – as far as we can tell from the selection of cases screened – reduced to tears. Moreover, in Culpeper (2005) I accounted for the fact that they took offence despite the presence of mitigating factors by noting a long line of research in social psychology showing that people are not very good at adequately factoring in context; targets of impoliteness are liable to be overwhelmed by the salience of impolite behaviour and not pay sufficient attention to potentially mitigating contextual factors (e.g. being targets of impoliteness is an expectable part of being a contestant on this kind of show).

My data consist of an interaction involving the contestant Susie. Performances on *Pop Idol* tend to be either very bad or very good. I have been informed by people familiar with auditioning for this show that the production team screen all contestants, and only let the bad and the good through to the actual audition in front of the judges. This makes good commercial sense, as it is the extremes of singing ability that are likely to prove the most entertaining. This does not mean that all the judges will agree on the quality of the performance. Evaluating singing is not an exact science. Also, it is clear from comments

made that the judges do not have a shared list of assessment criteria, instead deploying their own. All this leads to the possibility of conflicting views being expressed by the judges, and thus additional drama (of course, it is also possible that the judges are to some extent engineering conflicts for such dramatic purposes). Susie's performance triggers such a conflict, with two judges, Neil Fox and Nicki Chapman, expressing strongly positive evaluations, and another judge, Pete Waterman, expressing a strongly negative evaluation. This provides the opportunity within one piece of data to see how prosody works to express a positive evaluation and to express a negative evaluation, and moreover to see how those workings contrast (if they do). The full interaction is transcribed below (prosodic features are not marked, because relevant details will be provided in the discussion). The segments of texts in bold will be the focus of my analysis.

Transcription conventions: brackets indicate overlap, and the equals sign indicates 'latching' (change of speaker without an audible gap).

Judges N = Nicki Chapman; F = Neil Fox (alias 'Foxy'); P = Pete Waterman; S = Simon Cowell
Contestant SU = Susie

N: hi
SU: hi
S: how you doing
SU: I'm alright [yeah erm
S: [good erm your name is
SU: my name's Susie
S: and why are you here
SU: because this is my absolute craziest dream=
S: = yep=
SU: = in the history of the
      universe more than anything else I've ever wanted=
S: =yeah=
SU: =is this thing
S: that's it=
SU: =that's it
S: [ok
SU: [*indistinct*]
[*gives her performance*]
F: I think you looked great when you walked in
SU: thank you
F: you have lips every man wants to kiss
SU: [*laugh*]
F: but you've got a great voice
SU: oh thank you
F: I felt a bit of a tingle there
N: **you've got your own style, you've got your own look, lovely lovely
      voice**
SU: oh thank you
P: I must be hearing an apparition here, cos I don't see any of this
SU: oh
P: you said in the warm up it was your craziest dream
SU: it is
P: yeah that's what I think it is too a craziest dream, I don't see it
      at all, **it was an uninspired vocal, bored me, just nothing about it,
      it was lifeless**
SU: really
P: yeah
S: ok

SU: simon
S: I have to agree
SU: with Pete
S: no with the other two
SU: [*inhalation*] cos I always agree with what you think every time I've
       watched the show=
S: =no I love your voice=
SU: =every time
S: the great thing about this competition is two things number one we never
       agree and number two we offer it to the public, so let's find out now,
       Foxy
F: definitely must go through
SU: thank you
S: Nicki
N: yes
S: Pete
P: no
S: we're going to see you through to the next round congratulations
N: well done

5.2 Nicki on Susie

Nicki Chapman's evaluation, the emboldened segment in the transcription, is one of a series of compliments initiated by Neil Fox. Each compliment is followed by a receipt from Susie (e.g. "oh thank you"). All three seem to be doing 'politeness', as face is given by the compliment and by the gratitude expressed in the receipt. Figure 1 presents a visualisation of some of the prosodic features of Nicki's utterance. The figure consists of three tiers. The first at the top represents fluctuations in air pressure, providing an indication of relative loudness (intensity) and duration. The second represents changes in pitch (fundamental frequency expressed in Hertz) over time, providing an indication of the intonation contour of the utterance. The bottom tier contains the words that were spoken. You will see that the suggested intonation contour looks somewhat fragmented in places. This is not, of course, intervening silence (except where silence is recorded per second in the bottom tier), but simply that ambient noise prevented the computer from accurately recording the fundamental frequency. The reader needs to imagine the 'dots' joined.
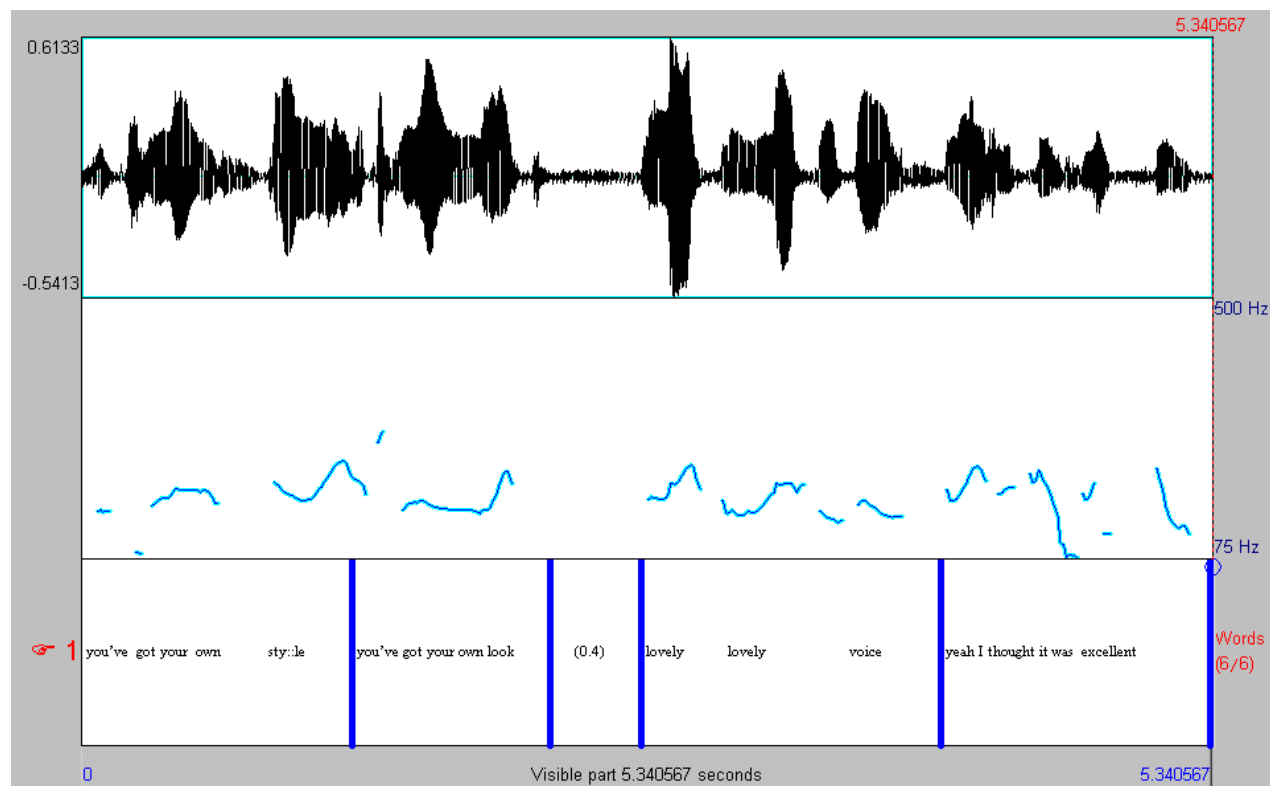
This is the final draft of Culpeper, Jonathan (2011) 'It's not what you said, it's how you said it!': Prosody and impoliteness. In: Linguistic Politeness Research Group (eds.) *Discursive Approaches to Politeness*. De Gruyter Mouton: Berlin, 57-83. It may contain minor errors and infelicities.

Figure 1. Nicki Chapman evaluating Susie

Perhaps the most notable feature of Nicki's evaluation is the prosodic parallelisms. The tone units "you've got your own style" and "you've got your own look" not only repeat grammatical and lexical structures, just varying the final lexical item, but repeat prosodic characteristics, each tone unit having a heavy accent and sharp fall on the final syllable. The segment "lovely lovely voice" contains what might be described as a triple down-step pattern: it has three similar contours, but each following peak lower than the previous. There is a natural tendency for pitch to decline generally during an utterance (a phenomenon labelled 'declination'), but this has been linked to rhetorical function, the most clearly defined triple down-step pattern being linked to the 'newness' of the information of the sentence (Wichmann 2000: chapter 5). Here, the rhetorical effect is more likely to be the emphasis of the loveliness of the voice. The final word of this segment, "excellent", has particularly marked pitch movement, compared with the other words, something which reinforces the positive evaluation expressed by that word. Finally, we can note the low onsets for all tone units.

Nicki's use of prosody is not specific to conveying a positive evaluation of being polite. In fact, the triple down-step pattern is used fairly often in the performance of impolite behaviours. Instead, prosody here clearly has a rhetorical function. As Knowles (1984: 227) puts it:
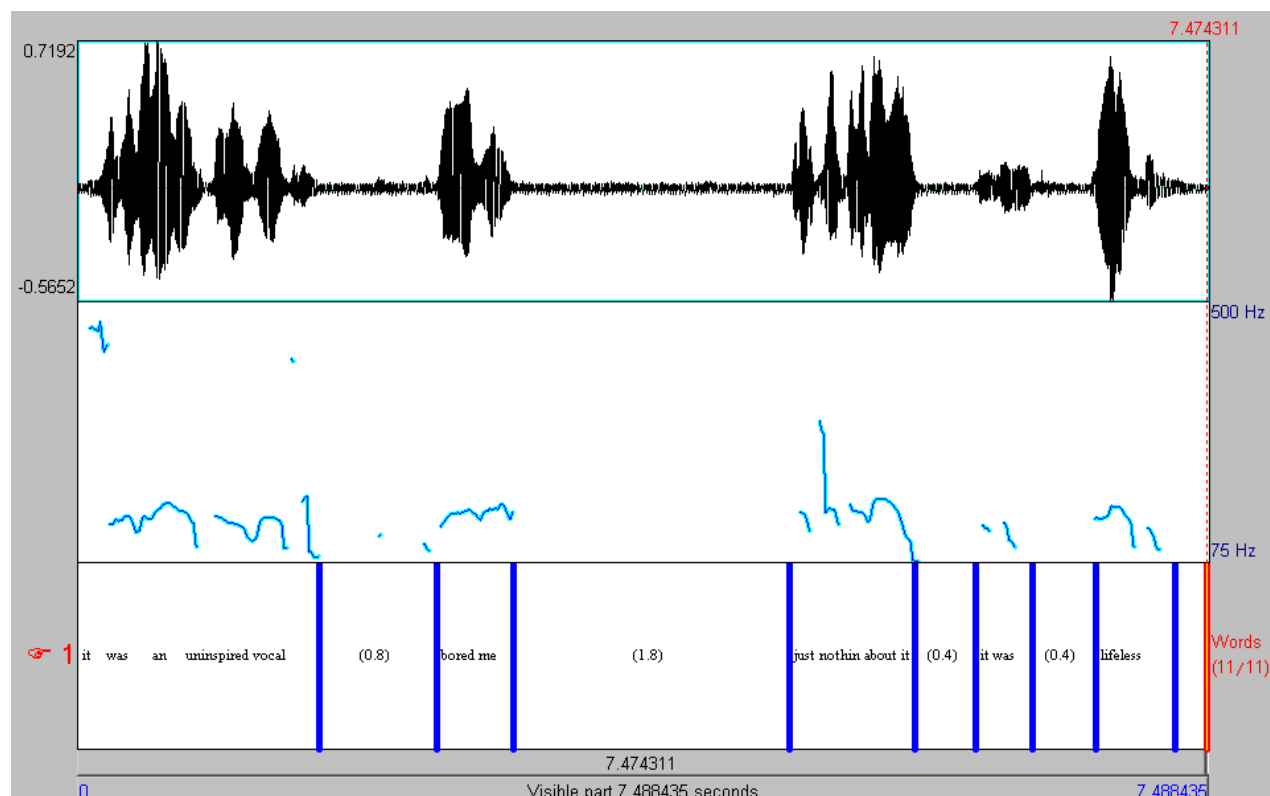
> The speaker has not only to decide what to say, but how to convey it effectively to the addressee. He [*sic*] has several channels at his disposal – verbal, intonational, paralinguistic – and employs communicative strategies to combine the signals sent on each channel so that the total effect will be correctly interpreted by the hearer. Conventional linguistics concentrates on the content of the message that is conveyed: intonation is part of rhetoric, or the strategies employed to get that message across.

We should also note the prosodic features I pointed out are all marked or foregrounded, against local norms constituted by other parts of the segment I analysed and also the prior utterances.

The important point about this polite interaction is that it creates a particular local norm, a high politeness threshold, that is violated in the following interaction, as we shall see.

5.3 Pete on Susie

Pete Waterman's disagreement with the other two judges' 'polite' positive evaluations is immediately flagged: "I must be hearing an apparition here, cos I don't see any of this". Figure 2 presents a visualisation of some of the prosodic features of Pete's utterance (emboldened in the transcription).

Figure 2. Pete Waterman evaluating Susie

The timing of Pete's speech is marked, compared with the local norms created by the previous speakers. He speaks 20 syllables in 7.5 seconds; in comparison, Nicki speaks 24 syllables in 5.3 seconds in the segment analysed above. One of the reasons it takes Pete longer is that he has more pauses, in fact four (of 0.8, 1.8, 0.4 and 0.4 seconds duration) compared with one (of 0.4 seconds duration). His speech in also rhythmically relatively monotonous: each tone unit ending in or consisting of a rise-fall on two syllables ("vocal", "bored me", "'bout it", "lifeless"). The segment "it was an uninspired vocal" has a very high pitched onset, but this is rapidly reset to relatively low and restricted pitch range. The pitch movement on "bored me" is very restricted, and his voice quality has some creak. Pete has a lower pitch average of 159.6Hz compared with Nicki's 173.7Hz. However, his baseline is likely to be different, given that he is a man. Nevertheless, even by his own local norm the pitch average of 141.4Hz for the final tone unit, "lifeless", is low. Also, it is interesting to observe that Pete has a much higher standard deviation in pitch values for his segment than Nicki, 68.0 against 32.1. Although Pete may be generally low in pitch, he incorporates some sweeping falls. The nucleus on the first syllable of "nothing" is made strikingly prominent through pitch height, and is followed by a sweeping fall. It is difficult to meaningfully measure Pete's loudness across this segment, because of the many pauses. Certainly, he seems quieter then Nicki, and the top tiers of their respective Figures display differing amounts of loudness, Pete's displaying less.

14

The prosodic features I have mentioned with regard to Pete's segment are marked against the local norm for Nicki: they reinforce Pete's disaffiliative stance. Further, Pete's prosody is not rhetorically supportive of his meaning in the same way as Nicki's. His prosody is not only marked but also indexical of the very meaning he expresses with his words – the performance bored him (cf. "bored me"). Some of his prosodic features – notably, low pitch range, slow speech, falling contours – are consistent with research on the acoustic correlates of boredom (see, for example, Scherer 1974; though the findings of research on boredom are not all clear-cut, cf. Scherer 1986: 161). Similar features have also been correlated with disgust. The correlates of disgust are said to include: very slow speech rate, much lower pitch average, slightly wider pitch range, quieter, grumbled, chest tone, wide falling terminal contours, normal articulation (Murray and Arnott 1993: 1104–5, 1106). Arndt and Janney (1987) would label this similarity of meaning cues in multiple modalities "redundant patterning" but, as I already pointed out, it is not redundant: it clarifies and amplifies the meaning he conveys.

6. Conclusion

I have emphasized the absence of study in the area of politeness or impoliteness and prosody. This chapter is a small step towards filling that gap and, more generally, raising awareness of the important role prosody plays in politeness/impoliteness. Assuming that readers are likely to be better versed in politeness or pragmatics than prosody, I have devoted some space to outlining what prosody is, what a prosodic description might consist of, the local and global issue, and the various sub-fields in prosody studies (i.e. linguistic-form, interpersonal, interactional). I stressed that prosodic features are gradient and relative to some degree, meaning that context must always be factored in when analysing the role of prosody in communication.

Politeness studies often fail to mention prosody at all. Brown and Levinson (1987) make a scattering of brief remarks, but it is clear that they are operating with a restricted model of communication. The single notable exception is Arndt and Janney (1987), which focuses on multi-modal emotive communication. They forge a link between increasing/decreasing emotional security and expressing politeness. Their detailed, systematic and replicable model treats prosody as a contextual phenomenon. However, their main focus is on modalities – looking at prosody that is attitudinally marked because one modality (e.g. prosody) contrasts with another (e.g. words and/or grammatical structures). Nowhere in their book do they tackle naturally-occurring conversation, and, in particular, how highly relevant local contexts are constructed.

My next step was to draw evidence from metapragmatic comments to support the idea that prosody plays a role – and possibly a key one – in the lay person's understandings of impoliteness. In fact, such comments revealed that particular prosodies (e.g. high pitched 'whines') in particular contexts (e.g. children making requests to parents) could evoke impoliteness. Also, it was clear that prosody was a useful way of conveying off-record impoliteness (Culpeper 2005) in public contexts where evoking impoliteness through on-record words and structures was not in the interests of the producer.

In the final section of the chapter, I examined the use of prosody in evoking impoliteness in its naturally-occurring context. More specifically, I carried out a detailed instrumental analysis of two segments from one interaction from the exploitative talent show *Pop Idol*. Prior to Pete's impolite contribution, a politeness context (repeated compliments with supporting prosody and their uptake) had been engineered in the previous discourse. Consequently, there is an expectation that the third judge, Pete, will be similarly complimentary. This locally created context is the most (psychologically) relevant 'norm' by which Pete's behaviour is marked; indeed, defeating this politeness expectation gives his impoliteness power. His prosody exacerbates the impoliteness already apparent in his verbal message by: (a) contrasting with Nicki's previous prosody (Pete has relatively slow speech, long pauses, monotonous rhythm, lower pitch range, wide sweeping falls mixed with restricted pitch movement, more voice creak); (b) indexing the attitude of boredom and disgust, and thereby creating an amplifying parallelism with

what he says in words; and (c) focussing attention on particular parts of the verbal message (e.g. prominent accent on "NOTHing about it").

You will have noted that my approach to prosody is somewhat eclectic. I do not, for example, simply espouse an interactional view. My analysis showed how prosody works, sometimes simultaneously, on several levels to contribute the evocation of impoliteness. The one particular area which I have drawn attention to is the role of the local prosodic context in the creation of impoliteness, something that seems to have escaped other studies.

Notes

1.   The final third of this paper on prosody was primarily authored by Anne Wichmann.
2.   Tony Cowell is Simon Cowell's older brother.

References

Archer, Dane and Robin M. Ackert
    1977   Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology* 35: 443–449.
Arndt, Horst and Richard Wayne Janney
    1985   Politeness revisited: cross-modal supportive strategies. *International Review of Applied Linguistics in Language Teaching* 23 (4): 281–300.
    1987   *Intergrammar: Toward an Integrative Model of Verbal, Prosodic and Kinesic Choices in Speech*. Berlin: De Gruyter.
Bargiela-Chiappini, Francesca
    2003   Face and politeness: New (insights) for old (concepts). *Journal of Pragmatics* 35 (10–11): 1453–1469.
Boersma, Paul and David Weenink
    2008   *Praat: doing phonetics by computer* (Version 5.0.40) [Computer program]. Retrieved November 10, 2008, from http://www.praat.org/
Brazil, David
    1985   *The Communicative Value of Intonation in English*. Birmingham: English Language Research/Bleak House Books.
Brown, Penelope and Stephen C. Levinson
    1987   *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
Cheang, Henry S. and Marc D. Pell
    2008   The sound of sarcasm. *Speech Communication* 50: 366–381.
Couper-Kuhlen, Elizabeth and Margret Selting (eds.)
    1996.   *Prosody in Conversation: Interactional Studies*. Cambridge: Cambridge University Press.
Cowell, Simon
    2004   *I Don't Mean To Be Rude, But…: The Truth about Fame, Fortune and My Life in Music*. London: Ebury Press.
Cowell, Tony
    2006   *I Hate to be Rude, but … Simon Cowell's Book of Nasty Comments*. London: Blake Publishing.
Crystal, David and Derek Davy
    1969   *Investigating English Style*. London: Deutsch.
Culpeper, Jonathan
    1996   Towards an anatomy of impoliteness. *Journal of Pragmatics* 25: 349-367.
    2005   Impoliteness and 'The Weakest Link'. *Journal of Politeness Research* 1 (1): 35–72.
    2007   Reflections on impoliteness, relational work and power. In *Impoliteness in Language: Studies on its Interplay with Power in Theory and in Practice*, Derek Bousfield and Miriam Locher (eds.). Berlin/ New York: Mouton de Gruyter, 17–44.
    Forthcoming *Impoliteness: Using Language to Cause Offence*. Cambridge: Cambridge University Press.

This is the final draft of Culpeper, Jonathan (2011) 'It's not what you said, it's how you said it!': Prosody and impoliteness. In: Linguistic Politeness Research Group (eds.) *Discursive Approaches to Politeness*. De Gruyter Mouton: Berlin, 57-83. It may contain minor errors and infelicities.

Culpeper, Jonathan, Derek Bousfield and Anne Wichmann
    2003   Impoliteness revisited: With special reference to dynamic and prosodic aspects. *Journal of Pragmatics* 35 (10–11): 1545–1579.
DePaulo, Bella M. and Howard S. Friedman
    1998   Nonverbal communication. In *The Handbook of Social Psychology*. Vol. 2, Daniel T. Gilbert, Susan T. Fiske and Gardner Lindzey (eds.) 4th edition, 3–40. New York: McGraw-Hill.
Eelen, Gino
    2001   *A Critique of Politeness Theories*. Manchester: St. Jerome.
Frick, Robert W.
    1985   Communicating emotion: The role of prosodic features. *Psychological Bulletin* 97 (4): 412–429.
Goffman, Erving
    1974   *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper and Row.
Gumperz, John
    1982   *Discourse Strategies*. Cambridge: Cambridge University Press.
Halliday, Michael A.K.
    1967   *Intonation and Grammar in British English*. The Hague: Mouton.
Knowles, Gerald
    1984   Variable strategies in intonation. In *Intonation, accent and rhythm. Studies in Discourse Phonology*, Dafydd Gibbon and Helmut Richter (eds), 226–242. Berlin/New York: Mouton de Gruyter.
Kruger, Justin, Nicholas Epley, Jason Parker and Zhi-Wen Ng
    2005   Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology* 89 (6): 925–936.
Ladd, Robert
    1996   *Intonational Phonology*. Cambridge: Cambridge University Press.
Locher, Miriam and Derek Bousfield
    2008   Introduction: Impoliteness and power in language. In *Impoliteness in Language: Studies on its Interplay with Power in Theory and in Practice*, Derek Bousfield and Miriam Locher (eds.), 1–13. Berlin/New York: Mouton de Gruyter.
Murray, Iain R. and John L. Arnott
    1993   Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America* 93 (2): 1097–1108.
O'Connor, J. D. and G. F. Arnold
    1973   Reprint. *Intonation of colloquial English*. Bristol: Longman. Original edition 1961.
Ofuka, Etsuko, J. Denis McKeown, Mitch G. Waterman and Peter J. Roach
    2000   Prosodic cues for rated politeness in Japanese speech. *Speech Communication* 32: 199–217.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik
    1985   *A Comprehensive Grammar of the English Language*. London: Longman.
Roach, Peter
    1991   *English Phonetics and Phonology: A Practical Course*. Cambridge: Cambridge University Press.
Sadanobu, Toshiyuki
    2004   A natural history of Japanese pressed voice. *Journal of the Phonetic Society of Japan* 8 (1): 29–44.
Scherer, Klaus R.
    1974   Acoustic concomitants of emotional dimensions: Judging affect from synthesised tone sequences. In: Shirley Weitz (ed.) *Nonverbal Communication*. New York: Oxford University Press.
    1986   Vocal affect expression: A review and a model for future research. *Psychological Bulletin* 99 (2): 143–165.
Shochi, Takaaki, Véronique Auberge and Albert Rilliard
    2007   Cross-listening of Japanese, English and French social affect: About universals, false friends and unknown attitudes. *Proceedings of the 16th Congress of Phonetic Sciences at Saarbrűcken, 6–10 August 2007*. Available at: http://www.icphs2007.de/conference/session16.html
Spencer-Oatey, Helen
    2002   Managing rapport in talk: Using rapport sensitive incidents to explore the motivational concerns underlying the management of relations. *Journal of Pragmatics* 34 (5): 529–545.
    2008   *Culturally Speaking: Managing Rapport through Talk across Cultures*. 2nd edition. London: Continuum.

Sperber, Dan and Deirdre Wilson
 1995   *Relevance: Communication and Cognition.* 2nd edition. Oxford, UK/Cambridge, USA.: Blackwell.
Terkourafi, Marina
 2007   Towards a unified theory of politeness, impoliteness, and rudeness. In *Impoliteness in Language: Studies on its Interplay with Power in Theory and in Practice*, Derek Bousfield and Miriam Locher (eds.), 45–74. Berlin/New York: Mouton de Gruyter.
Trager, George L.
 1958   Paralanguage: A first approximation. *Studies in Linguistics* 13: 1–12.
Watts, Richard J.
 2003   *Politeness*. Cambridge: Cambridge University Press.
 2008   Rudeness, conceptual blending theory and relational work. *Journal of Politeness Research* 4 (2): 289–317.
Wichmann, Anne
 2000   *Intonation in Text and Discourse: Beginnings, Middles and Ends*. Harlow: (Longman) Pearson Education.
Zuraidah, Mohd D. and Gerry Knowles
 2006   Prosody and turn-taking in Malay broadcast interviews. *Journal of Pragmatics* 38 (4): 490–512.