

Prosody and Particles: A Study of Interaction in a Malaysian Academic Meeting

ZURAIDAH MOHD DON
Faculty of Languages and Linguistics
University of Malaya, Malaysia
zuraida@um.edu.my

TAM SHU SIM
Faculty of Languages and Linguistics
University of Malaya, Malaysia

ABSTRACT

This article examines prosody in interaction in the context of a 66 minute meeting to agree on assessment marks as part of a pilot project studying assessment procedures. The data is drawn from a larger set of data of approximately 10 hours of academic meetings. The three participants are female colleagues involved in English Language teaching. The raw data in the audio file was subjected to a three stage process of annotation, analysis and presentation. Findings show that particles can be sorted into emotive particles, artefacts of vocalisation, and those concerned with the sharing of knowledge between speaker and addressee. They can also be classified according to their relationship to speech, and whether or not they are linguistically and phonologically organised. What is also interesting is that when prosody co-occurs with particles it operates in a different way than when it illustrates syntactic or information structure or plays a role in turn taking. Prosody appears to operate as an independent component of particles, using salience to draw attention, or lack of salience to accompany activities which are not intended as interruptions.

Keywords: prosody; academic meeting; particles; English language; interaction

INTRODUCTION

Prosody straddles the boundary between linguistic and non-linguistic phenomena. In the study of prosody there has always been a minority interest in going beyond linguistic structure (Bolinger 1972, 1986, 1989; Barth-Weingarten, Dehé & Wichmann, 2009 for the contribution of prosody to meaning). Kingdon (1958) and Crystal (1969) presented a wide range of prosodic phenomena going far beyond linguistic information, and O'Connor and Arnold (1973) sought to trace in particular the connection between prosody and attitude (Monetta, Cheang & Pell 2013). Even in the case of prepared speech (Knowles, Williams & Taylor 1996), although the prosody is undoubtedly largely concerned with reflecting syntactic and information structure, there is nevertheless a residue that at the present state of knowledge is difficult or impossible to explain. This residue includes a number of particles such as [ə:] and [m:], some of which are clearly not part of language. Others, such as *right*, appear to have been coopted from language for a different communicative purpose, and retain some of their linguistic properties. These particles have been chosen for study here because they make a useful link between the familiar territory of language and *terra incognita* beyond.

This paper is concerned with the place of prosody in interaction, taking ideas as appropriate from Conversation Analysis (CA), and concentrating on the prosody associated with particles popularly known as *ums* and *ahs* (Lam 2009). These particles are examined in the context of a 66 minute academic meeting to agree on assessment marks as part of a pilot project studying assessment procedures (see also Ting, Mahanita Mahadhir & Chang 2010 focusing on grammatical inaccuracy in oral communication assessment). The data is drawn from a larger set of data of approximately 10 hours of academic meetings. The analysis put

forward in the section entitled Results is based on about 66 minutes of talk. The three participants are female colleagues involved in English Language teaching. The aim of the research reported here is to go beyond the study of prosody in relation to turn taking on the one hand, and linguistic structure on the other. We address two questions: What particles are the typical responses that mark the end of turn-constructive units (TCUs), and what are their phonetic form and the associated prosody. TCUs are the basic units of conversation, each ending in a transition relevance place (TRP), that is, the place where the turn may shift to another speaker (Sacks et al. 1974).

LINGUISTICALLY UNORGANISED PARTICLES

The particles under investigation are referred to in various ways including ‘discourse markers’ (Schiffrin 1987), ‘discourse particles’ (Aijmer & Simon-Vandenberg 2003) or ‘interjections’ (Enfield et al. 2013), in addition to the expression *ums* and *ahs*. They differ from words in having little or no linguistic organisation. This means *inter alia* that they have no grammar¹. Instead of being included in syntactic structures, particles are typically preposed or appended to structures, located in pauses between structures, or uttered simultaneously with another speaker’s structures (Fraser 1999). In certain circumstances, they can however fill syntactic slots, e.g. *um* and *ah* function as nouns in the phrase *ums and ahs*, and *ok* functions as an adjective in *I’m a lumberjack and I’m okay*. They do not belong to the semantic structure of the language, even if they look like words, e.g. *Right?* as a particle has little connection with its syntax or meanings in *right answer* or *human rights*. The “meaning” typically relates to the ongoing situation, which in many cases corresponds to an identifiable stage in a script (Schank & Abelson 1977). Particles have phonetic form, since they are produced by normal phonation and articulation, but their phonological status is variable. Particles of the type [ə] and [m] resemble isolated phonemes, while *erm* and *ya* resemble syllables. Some words such as *well* (Aijmer & Simon-Vandenberg 2003) and *right* when coopted for use as particles retain their phonological structure, but others lack any clearly defined phonological organisation, which means that they are not subject to phonological rules

If these particles are not linguistically structured, we cannot assume that they can be analysed using linguistic techniques. The first attempt to describe English prosody (Steele 1775) actually took a musical approach. Twentieth century work was influenced by Armstrong and Ward (1926), who described English intonation in terms of two ‘tunes’ and their variations. A richer set of tunes was described by O’Connor and Arnold (1973), who attempted to relate different tunes to different attitudes. The first American approach to English prosody (Pike 1945, Wells 1945) sought to treat different pitch levels as phonemes, and so bring prosody within the ambit of linguistic theory; but this approach was ultimately unsuccessful. Prosody is more effectively represented using some more musical notation scheme such as British tonetic stress marks (Kingdon 1958, Knowles et al. 1996, Cruttenden 1986) or the tones and break indices (ToBI) system based on the work of Pierrehumbert (1980).

It is significant that all the prosodic patterns described by Reed (2011) to handle prosody in conversation can be better described as properties of music, treating prosody as a derivative of music. In comparison with music, prosody has no proper pitch scale, irregular and often ungrouped beats, and ill-defined dynamics. The musical approach has in recent decades received support from neurological studies (Falk 2000), for although language is generally claimed to be processed in the left hemisphere (for typical right-handed

individuals), pitch and the emotional content of speech and music are associated with the right hemisphere.

METHOD

APPROACH AND DATA

The approach is based on the British tradition of analysing prosody, and extends it using ideas from CA to transcribe and analyse recordings of naturally occurring data (Heritage & Clayman 2010). The focus is on discourse particles and pitch movements are annotated where possible using the British tradition of tonetic stress marks (Kingdon 1958, O'Connor & Arnold 1973). The data consists of an audio recording lasting about 66 minutes of a meeting of three female academic colleagues to agree assessment marks as part of a pilot project studying assessment procedures. It is drawn from a corpus of approximately ten hours of talk. The participants are C, who is the chair, the head of the department H, and a language teacher T. All three colleagues are involved in English Language teaching, but each has a first language other than English. The raw data in the audio file was subjected to a three stage process of annotation, analysis and presentation.

ANNOTATION

The annotation was carried out manually, every annotation label corresponding to an exact location in the waveform. The first approximation to phrase level annotation was made using normal spelling and marking the beginnings and ends of phrases. This procedure automatically measured pauses, including pauses between turns, and also overlap (where the pause duration < 0).

Standard spellings were used to represent words with predictable pronunciation, phonetic transcription being used to represent unpredictable detail, e.g. 'ye[s:]' represents the word *yes* with a lengthened final /s/. While T used a kind of prosody that could be represented by conventional transcription as for British English (Knowles et al. 1996), C and H used a different system with a similar role in reflecting syntax and – in part – information structure². Conventional punctuation with fixed values was used to represent pitch movements for these two speakers: a comma represents a low rise, a semi-colon a higher (and often crescendo) rise into the middle range, and a full stop represents a fall into the lowered range. A single exclamation mark '!' represents a fall from high pitch, and a question mark '?' the high rise. Combinations are also possible, e.g. '?!' represents a high rise fall. Another occasional combination is the low fall rise, marked '.,'. For all participants, the same transcription system was used for propositional content and particles (see Appendix 1 for the transcription symbols).

While annotations are motivated by observable events in the waveform, this is not always straight forward. In the case of pitch movements, because F0 contours are often interrupted by voiceless segments, or stretches of creaky or breathy voice. In some cases the direction of movement is identified from the trajectory of movement.

Since it was not evident what kinds of patterns would be found, observations were carried out on everything that might be relevant. Voice quality emerged as probably relevant, and so five voice qualities were represented – modal, nasal, breathy and creaky voice and falsetto – as consistently as possible. The beginning and end of each departure from modal voice was marked in the annotation file, and in the final transcript, voice qualities were marked as follows:

Breathy voice: ‘. .’ Quiet voice: ‘◦’
Creaky voice: ‘˘’ Nasal voice: ‘_N’

The doubled exclamation symbol ‘!...!’ marks the beginning and end of speech using an ingressive air stream, and ‘_L’ of laughing.

Pitch range was divided into three sub-ranges, upper, middle and lower. It was then observed that voice qualities were associated with positions in the overall pitch range, so that there could be recurrent patterns combining voice quality and pitch range. In every case, the prosodic pattern accompanied a phonetic event that could be handled by conventional phonetic transcription.

At the same time, notes were made on the function of the particles. At first, these were *ad hoc* observations relating to the situation in which individual items occurred, e.g. as the answer to a question. The comparison of large numbers of examples led to the identification of patterns, e.g. *ok* was clearly related in this set of data to moving on to the next item on the meeting agenda. As different types emerged, the transcription became less hit-and-miss and more systematic. The particle types are described in the results section below.

ANALYSIS

In the second stage, the contents of the annotation file were written to a table in a Microsoft Access database. The analysis was then based on the contents of the table, with reference wherever necessary back to the waveform. Annotation cannot in practice be done once for all, and in the course of the analysis it was necessary many times to return to the waveform to extract more information or to improve the accuracy of earlier annotation labels. In this way the entire analysis was based on information explicitly recorded in the annotation file. And since the annotations represent acoustic events in the waveform, the analysis actually deals with events taking place in the real world.

When making the analysis, it is essential as Local points out (1996, p. 179) that “the categories of the analysis are carefully warranted, or justified, by the interactional behaviour of the participants themselves, and not simply by the armchair intuitions of the analyst”. In our data, there is in many cases no audible response at all, even to items that seem to require a response. Some additional means of analysis is clearly required. There is always a danger that the examination of a single item, or a small set of items, no matter how carefully the situational context is investigated, will lead to an *ad hoc* explanation of some phenomenon which does not apply beyond the immediate context. This danger is progressively reduced as more and more data is included in the analysis, and recurring patterns begin to emerge. Recurring patterns are less likely to emerge by chance, and explanations which apply to large sets of data are better founded. In this case, the audio recording was treated as a small corpus, and transcribed and analysed systematically and exhaustively. Figures are given below to indicate the size of the dataset on which the analysis is based, and the fragments presented as examples are chosen as good illustrative examples of their type.

THE TRANSCRIPT

The third stage was to create the transcript, which combines CA conventions adapted from Jefferson (2004; see Appendix 1), with phonetic symbols as necessary. The table contents were first formatted and then copied into an html file, which has the advantage that it can be opened directly in Word. All phonetic characters were represented by their Unicode values, and times were rounded off to the nearest centisecond. This procedure makes a logical

distinction between the nature and presentation of data. For example, schwa is represented by ‘@’ in the annotation file, but it is replaced by ‘ə’ in the transcript in accordance with the transcription convention in phonetics; and it could equally well be presented as ‘uh’ if that were for some reason to be preferred. Similarly, pauses of less than 0.20 seconds could if required be represented by the symbol ‘(.)’ following transcription conventions in CA.

The extracts from the transcript presented here contain all the relevant information transcribed. Unpredictable aspects of pronunciation are indicated by phonetic symbols in square brackets. The intonation used by speaker T was transcribed using British tonetic stress marks, and a system using conventional punctuation marks with fixed values was specially devised to transcribe the intonation used by speakers H and C.

The measurement of pauses to the nearest centisecond means that latching has a value of 0, and overlap less than 0. The negative values used to indicate overlap have the advantages of objectivity and logical independence from the manner of presentation on the printed page. Other transcription conventions are explained at the appropriate point in the sections below.

A noteworthy aspect of the audio recording was the range of voice qualities used. Voice quality was accordingly marked in some detail in the transcription, concentrating on falsetto in addition to modal, nasal, breathy and creaky voice. The preliminary analysis, however, did not bring to light any independent function for voice quality. Breathiness and creakiness tend to accompany low pitch, particularly a low fall in pitch, and to have much the same distribution as quiet modal voice. Quiet voice, breathy voice and creaky voice were accordingly grouped together as variant manifestations of what might be called *lowered voice*. When voices are lowered, they also tend to become breathy or creaky.

Lowered voice proved to possess interesting properties. In the normal case, we can expect to find evidence in the waveform consistent with the transcription. But if the voice is breathy or creaky, there may be little or no evidence of laryngeal activity, and so no identifiable fundamental frequency. In the absence of clear formants, it is not always obvious what the transcriber is responding to when identifying an instance of *ya* or [m:]. In some cases the observable laryngeal activity seems to indicate that the voice is as it were dipping into the lowered range; in such cases it could be a matter of chance whether laryngeal activity is sufficient to give the impression of pitch on the downward or the upward slope, and so whether the item is judged to have a fall or a rise in pitch. In line with conventional practice, particles were tentatively transcribed with pitch movements, but this is a matter that requires further research.

Similarly, falsetto seemed to occur in predictable circumstances, at the end of a high rise in pitch, or simultaneously with a pitch reset at the beginning of a new topic. The symbol ‘↑’ was used to mark the beginning and end of high voice, possibly including falsetto. An apparent exception to the general pattern is the frequent use of creaky voice to accompany the high rise on the particle *right?* In fact, all three speakers tend to use creaky voice on the vowel preceding a final voiceless stop, and so this is fully accounted for at the level of allophonic variation. Creaky voice is a predictable property of the vowel of *right*, and so does not constitute an independently meaningful prosodic pattern. The middle range between high voice and lowered voice is the default, and it is associated with modal voice. It is sometimes used with creaky or nasal voice, but the function of these, if any, has yet to be ascertained.

RESULTS

PARTICLES MARKING THE END OF TCUs

Since linguistically unorganised particles do not belong to a semantic system, it is difficult to specify their meaning. A possible starting point is the observation in paragraph 43 of Wittgenstein (1953) that “the meaning of a word is its use in the language”. *Mutatis mutandis*, the meaning of a particle is its use in a system of communication, this use being measured by its distribution. It is sometimes argued or implied (Guthrie 1997) that particles used interchangeably in the same context are alike in meaning; but here each particle is assessed in its own distribution. The first task is to identify recurring contexts which illustrate the distribution of different particles.

The end of a TCU is typically accompanied in this set of data not by competition for the floor, but by responses in the form of particles. According to Yngve (1970) these backchannel signals or turn-continuers (Schegloff 1981) indicate that the speakers do not want to take the floor. The responses will often be elicited by means of particles with a high rise in pitch, e.g. *You know?* or *Right?* Table 1 gives frequency counts with the eliciting particles in the columns and the responses in the rows. In every case the response has a low fall in pitch.

TABLE 1. Responses to high rises

	You know?	Right?	[ə]?	Ya?	YNQ	Ok?	[a:]?	Alright?	[m:]?	Is it?	Total
zero	27	16	17	8		6	4	4	1		83
ya	7	7	3	11	2				1	1	32
yes		2	3	2							7
yup		2									2
yeah			1	1							2
ya lah							1				1
no					2						2
[m:]	1	4	3	3			1	1	1		14
[mhm]			1								1
ok	1				1						2
alright			1								1
right				1							1
text		4			4						8
Total	36	35	29	26	9	6	6	5	3	1	156

The table counts instances in which one particle follows another. This does not state whether the second particle links back to the first particle, or whether it links forward to new material (Enfield et al. 2013, pp. 345-347). There could logically also be the Janus type, linking both ways simultaneously, and the type that links to neither and just places a boundary between them. The most frequent type is in fact zero, which leaves the observer to infer that there is a boundary at this point.

The other frequent type is *ya* and its variants including *yes*, which intuitively seems to treat the high rise as an indicator of some kind of question, to which it is an appropriate response. The modal response to *Ya?* is *ya*. The status of [m:] remains unclear at this stage.

The second context considered is a medial pause in a TCU, to which the addressee responds by using a particle as a pause filler. Frequency counts are presented in Table 2. In this case several different pitch movements are used, and so the particles are in the columns and pitch movements in the rows.

TABLE 2. Responses to mid-TCU pauses

	[m:]	[mhm]	ya	Yes	yeah	[a:]	Ok	Total
Low fall	15	4	9	3	1	1	2	35
High fall			2					2
High rise	1							1
Low rise	1	2	1				7	11
Total	17	6	12	3	1	1	9	49

In this case the modal response is [m:] and its variant [mhm]. The figure for [mhm] includes its own variant [əhə], which is basically the same event, but with the mouth open, so that the air stream flows out of the mouth instead of being diverted through the nose. Although the frequency of *ya* and variants is lower, it is not far behind. Perhaps [m:] is a neutral pause filler, while *ya* is more positive.

Particles are also used at the beginning of a new TCU. In Table 3, the particles are again in the columns and pitch movements in the rows.

TABLE 3. Beginning a new TCU

	ok	Er	erm	Alright	ya	Oh	[a:]	Total
Low fall	12	1		1	1		1	16
High fall	1					2		3
Low rise	8							8
Mid rise	5							5
High rise	5							5
Level		3	6					9
Total	31	4	6	1	1	2	1	46

The most frequent type here is clearly *ok*. The low fall indicates that it is in the stand-alone position, before the beginning of the new TCU proper, whereas the low and mid rises link it to the following TCU. The high rise *ok?* also stands alone, and is equivalent to something like ‘Can we move on?’ A new TCU can also begin with a hesitation pause filler, written ‘er’ or ‘erm’. These, as already noted, are unusual in having an unchanging F0 contour. *Oh* is also used here, but infrequently. The last context considered concerns a TRP not embellished by a particle with a high rise, but which elicits a particle from the addressee. The frequency count is presented in Table 4.

TABLE 4. Responses at a TRP

	ya	Yes	yeah	[m:]	[əhə]	ok	oh	[a:]	Total
Low fall	27	5	1	21		8	2	2	66
High fall	1			4			1		6
Low rise	1				1	10			12
Mid rise						2			2
Total	29	5	1	25	1	20	3	2	86

Again, the modal pitch movement is a low fall. *Ya* and its variants are again closely followed by [m:] and its variant [əhə].

PARTICLES AND PITCH

Particles would appear to combine two independently meaningful components, namely phonetic form on the one hand, and the combination of pitch and voice quality on the other. High voice³ is by nature salient, and has an obvious function as an attention device akin to pointing (Tomasello 2010), as though the speaker is pointing with the voice. The high rise is well known as the ‘questioning’ tone, and points to an item that is worthy of attention. Lowered voice is associated with what Yngve (1970) called the back channel, and it essentially does not draw the attention of other participants. Items produced on a lowered

voice are not regarded as interruptions, although they can and often do count as answers to questions and responses to high rises. They are transcribed with a low fall by default, although instances transcribed with a low rise are also found in the transcript.

THE INVENTORY OF PARTICLES

At this stage, we can make some tentative suggestions concerning the inventory of particles used in this set of data. Since they do not constitute a semantic system, there are no clearly defined contrasts, and their boundaries are even fuzzier than those of lexical items. This section complements the previous section by examining each particle in turn in its different contexts.

This section begins with the three expressions *alright*, *right* and *you know*, which being coopted from the language system for use as particles can reasonably be expected to retain something of their lexical meaning. These are then used as models for the remaining particles. The numbers in brackets refer to the number of instances used for the analysis.

ALRIGHT (16)

Like the linguistic expression *all right*, *alright* is used to check that the situation is satisfactory or acceptable. *Alright?* with a high rise (5) can be paraphrased ‘Is the situation acceptable?’ as exemplified in 3C.

[3238.14] 1C (0.28) That means; part of, the instruction; is they can

[3239.88] 2H (0.15) ◦[m:].◦

[3239.90] 3C (-0.43) play_back. (0.52) .ok.. (0.41) alright?

[Key: The figure in the leftmost column in square brackets gives the time in seconds relative to the beginning of the file. H and C (and elsewhere T) represent the speaker. The figure that follows immediately in brackets represents the time lag from the previous end of talking. Figures inserted into continuous text represent the duration of pauses.]

Here, H fills C’s pause before C completes her decision with a falling *ok*. She then seeks confirmation that her decision is acceptable. With a low fall (9; see 3C), the meaning can be paraphrased ‘(I take it that) the situation is acceptable’.

[62.06] 1C (-0.05) yeah. (0.09) use, the second, set.

[63.14] 2H (0.00) ok! (1.00) alright.

[64.58] 3C (-0.41) ok. (0.16) al ight.◦

Here *ok* marks the end of decision making, and *alright* the acceptability of the decision.

There is one case of *alright* (3T) with a low rise:

[3266.72] 1C (-0.18) [nu:l]⁴ do <that. [ə]?

[3267.40] 2H (-0.11) [m:].

[3267.67] 3T (-0.09) alright,

[3267.99] 4H (-0.02) ◦[m:].◦

[Key: The symbol ‘<’ marks an accented function word.]

But this example is difficult to interpret since the whole context is incoherent. This and other cases of incoherence are probably more worrying for the analyst than for the original participants.

The remaining case *alright!!* (2H) with a very high fall appears to be unrelated in meaning, and can be paraphrased by something like ‘cease and desist!’

[3083.39] 1C (0.36) earlier; I wasn’t, allowed, to speak.

[3083.56] 2H (-1.11) Alright!! (0.48) yes yes yes.

RIGHT (44)

Where *alright* is about the ongoing situation, *right* appears to be about shared knowledge. The use of a high rise *right* by speaker 3H indicates that she positions herself as having the knowledge (for the concept of *epistemic status* see Heritage, 2012), even though the addressee C might be the primary owner of the information in question, and checks to ensure that the addressee also knows.

[0.25] 1H (0.25) we now,

[0.53] 2C (-0.29) (o)k.

[0.82] 3H (0.00) have to decide. right?

[1.64] 4C (0.00) ya.

C as chairwoman is the primary owner of the knowledge of the purpose of the meeting, but it is made explicit instead by H.

With a low fall (5), *right* can be used to initiate a new topic, in this case after some hesitation:

[3239.90] 1C (-0.43) play_back. (0.52) ..ok... (0.41) alright? (0.18) ...a[n:]d (2.12) ![h]! [ə]m right.
>now today; the combined, marksheet; was not ...

[Key: the symbol “>” marks a word that might have been accented, but is left unaccented.]

Or to confirm agreement (4C)

[3424.68] 1H (0.05) So certain, format.

[3426.10] 2C (0.18) ya.

[3426.19] 3H (-0.25) · ·might· · come, in.

[3427.28] 4C (0.48) right.

The one case of a high fall confirms agreement already expressed by *ya!* (2C):

[3311.09] 1T (0.16) [kəz:] (0.08) a ˘lot of ·things have ˘happened since ˘then˘.

[3313.99] 2C (0.10) ya!

[3314.36] 3T (-0.19) ˘hahaha˘

[3314.74] 4C (-1.08) right!

There is no single meaning associated with the particle *right*. Meaning and use are associated with phonetic form and prosody together as units.

YOU KNOW (57)

You know refers to knowledge which is assumed to be shared, but which needs checking. With the high rise (35), it checks that knowledge is indeed shared:

[3699.59] 1C (0.29) °ya.° (0.27) O,k. (0.59) ₙ[ə:m]ₙ (0.17) ↑I think, that’s,↑ about <it. you know?

Here, C checks that all participants agree that the agenda has been completed. With a mid rise (15) on ‘you know’, it is more of a warning:

[2527.58] 2H (0.06) It’s like too, much. to process. you know;

With a low rise (7) on ‘you know’, the knowledge is known to be shared and no check is necessary:

[2876.78] 3H ↑how would you, ↑ stre[s:]. and (0.44) you know, intonation. and and make it, like lively.

C and T can be assumed to be familiar with the term *intonation*.

[m:] (184)

[m:] is the particle used by default by the addressee to fill the speaker’s pauses, and it is produced in most cases with a low fall (164), although three of these, all produced by H, could also be transcribed as low monotones. It is the counterpart to *er* [ə:] and *erm* [ə:m], which speakers use to fill their own pauses, but it is produced throughout with a closed mouth which diverts the air stream through the nose. If it has a meaning, it is something like ‘message received’. The use of a high fall (12) points to something significant in what the speaker has just said, and may indicate strong agreement or even approval⁵. A high rise (8) suggests a query about what the speaker has said, including ‘have you received my message?’.

YA (262)

In a positive context, [m:] can be assumed to imply some kind of agreement. *Ya* [jə] or [jə:] and its variants is explicit, and not only gives positive answers to yes/no questions, but also generally positive feedback. It would be rather odd or hypocritical for our participants to respond with *ya* while seriously disagreeing with what the speaker is saying.

Ya usually has a low fall (221), but a high fall also occurs (10), one of which is actually a high rise fall, and possibly signalling strong agreement. A further six instances are transcribed with low rises, but these could also be interpreted as low level tones. The high rise (34) suggests the paraphrase ‘do you agree with me?’ and often elicits the response *ya* with a low fall.

Variants of *ya* pattern in similar ways. *Yup* [jəp] (17) is essentially the same event but cut off as the mouth closes, and usually has a low fall (16), but there is one case of a doubtful low rise. *Yes* (53) has a low fall except in one case of *yes!* and a mid rise introducing one case of indirect reported speech. *Yeah* [jɛ:] (11) has a low fall except in one case where it has a high fall.

OK (125)

In other contexts, *ok* has a variety of functions including completing a section of talk involving a request or a suggestion. But in this particular set of data, *ok* seems to be concerned with moving on to the next item of business, as shown in Table 3. The meeting is highly structured, each candidate being considered in turn, and each component of the mark being considered in turn before an overall mark for the candidate is agreed. Since the participants share a detailed knowledge of the script, or the different stages which the

meeting has to go through, it is understandable that they will use particles in connection with moving on to the next item of business. The most important item is *ok*.

Ok with a low fall (60) generally terminates previous business, but as in this example, it may also be of the Janus type:

- [150.28] 1H (0.40) Was she clear, to you?
[152.50] 2T (1.20) [ə:m] (1.88) ˜not ˘very
[156.78] 3H (0.28) ə[m:] (0.09) ok. So (0.33) so because, of little, output;...

Here the freestanding *ok* with a low fall simultaneously marks the completion of the question-answer pair and prepares for the next point to be made.

The high fall (3) draws attention to moving on. In this next example, C calls the meeting to order:

- [3664.44] 1T (-0.99) ˌhahahaː
[3664.45] 2C (-0.95) ok! (0.29) whe[n:] (0.08) whe[n:] the padi; fields; are ripening;

Ok with a mid rise (9) introduces a new topic:

- [187.17] 1H (-0.06) ˆvery, experienced. ˆ
[188.28] 2T (0.19) ˌhaː
[189.20] 3H (0.60) ok; modest. right?

Whereas a low rise (44) continues the same topic:

- [1123.80] 3H (-1.44) ok, let's let's explore where we differ.

With a high rise (9), *ok* checks the appropriateness of moving on, i.e. 'Can we move on?'

- [3062.91] 1C (0.04) ok? (0.53) nothing more, to add?

The next example is slightly unusual, in that H utters *ok?* before completing the current item and proceeding to the next:

- [2448.78] 1H (0.03) ya. ok? (0.89) so that means, the second; and the fourth; (0.14) tend, to be at the same; (0.85) The next; is the third, student;

If this analysis of *ok* is correct, then it is essential to know what prosody is being used in order to evaluate it. Otherwise one has to invent the prosody for oneself (see, for example, the extracts in Beach 1993).

OH (16)

Oh is associated with the receipt of new information, and has been the topic of considerable previous work (Heritage 1984, Local 1996, Schiffrin 1987). It has an unusual distribution because it most frequently occurs in our data not only with a low fall (3) but also a high fall (6) and even a rise-fall (3).

First the low fall:

[417.19] 1T (1.45) n[ə:]N (0.12) eight.

[420.36] 2H (2.04) Fulfills task competent. oh. you put her eight.

H could have responded to the mark with [m:], but instead cites the regulations. She utters *oh* with a low fall when she realises she has been given the mark.

This next use of the high fall indicates that the new information does not have to come from another person:

[3268.32] 1C (0.05) ok? (0.81) So (0.18) oh! (0.14) ↑Did you find; ↑ that the notes were, very important to, jog, your memory?

C recalls a question she was intending to ask but must have temporarily forgotten.

The rise-fall is sometimes associated with information received which is contrary to expectations:

[390.89] 1T (0.12) [kç] .[h].. ɿ rated her the ʌhighest | with ɿtwenty ɿnine ||

[395.44] 2H (0.78) Oh?!

H here responds to a figure outside the expected range. She has to delete the expected figure and replace it with the new figure.

Oh is also used to introduce – ‘prefaced’ to (Heritage 1998) – a particle or phrase. Examples in the data include one each of *oh ya.*, *oh ya!*, *oh no.* and *oh ok.* *Oh* has to come first, so that expressions such as **ya oh* are ill formed. This means that these expressions have a rudimentary syntax, although not the complex syntax associated with language.

DISCUSSION

This discussion section moves on from the description and analysis of particles to two important sets of generalisations that can be made about their use in a wider context. Where appropriate, examples from the data are used to illustrate and justify the generalisations and relevant works are cited to support the argument.

THE HUMAN VOCAL REPERTOIRE

Vocal activity is conventionally viewed from the perspective of language, which puts phenomena such as hesitation, audible breathing and particles at ill-defined locations somewhere on the periphery. To see more clearly how these relate to each other, we have to adopt a different perspective, and treat them as components of the human vocal repertoire.

The particles with which this paper has been mainly concerned are connected with what Premack and Woodruff (1978) called a theory of mind, which involves the capacity to access what is going on in the minds of others. Although it is not essentially linguistic in nature, it is the enabling condition for Grice’s (1975) maxim of quantity, and for the construction of recursive sentences such as *I think you know he’s lying*. It is closely related to the CA concept of epistemic status (Heritage 2012, 2013).

Much of the work done by particles is concerned with harmonising the shared knowledge of speaker and addressee. This goes beyond the concept of *parity* (Lieberman 1996, pp. 31-32), which refers to a condition in which the message received is formally the same as the message transmitted by the sender. But transmitting linguistic information is only part of what humans do in communication. It is also important to ensure that the inferences drawn by the addressee and the understandings constructed match those intended by the

speaker. High voice would appear to be concerned with the assertion or querying of knowledge to be shared. A high fall is used in a situation in which either speaker or addressee has to modify the existing state of knowledge. A high rise which typically begins in the middle range ends in high pitch and possibly falsetto interrupts the flow of talk to elicit a response from the addressee to confirm or ensure that knowledge is shared. Eliciting a response necessarily means handing over the turn, but the speaker typically resumes the turn after a brief response from the addressee. In our data the response is often in lowered voice, indicating that this question-answer pair is incidental to the main turn taking.

PARTICIPANT RELATIONS

The meeting that contributed our data is all about sharing knowledge, and was accordingly rich in particles, many of which were concerned with knowledge sharing. It formally achieves its purpose when the participants agree on a set of marks, but from a human point of view it is important that they do so in a congenial manner, without falling out and becoming enemies. In this regard, we get a very different impression of the temper of the meeting depending on whether we look just at the words, or consider the interaction in a wider context.

C is the senior colleague in years of service, but she is a member of H's department and does not hold an academic post. It is therefore not clear who should lead the meeting, the chairwoman or the head of department. The meeting begins with a contest for control, which is won by C in the short term:

[0.25] 1H (0.25) we now,
 [0.53] 2C (-0.29) (o)k.
 [0.82] 3H (0.00) have to decide. right?
 [1.64] 3C (0.00) ya. (0.69) so >good afternoon, ladies; this is the last >part, of our piloting;

H starts to speak, but C overlaps with *ok* to start the meeting, and while she replies *ya* to H's question, she then takes over a minute to set out the procedures for the meeting. H then takes over, and much of the meeting consists of her discussion of marks with T. After some fifty minutes, C initiates a discussion:

[3073.01] 1C (0.81) we will now, move into, a discussion. (1.56) shall we?
 [3077.27] 2H (0.58) o. k... (0.28) I thought, we <did.
 [3079.05] 3C (0.66) ..[m:]?..
 [3080.18] 3H (0.93) ↑What is, the discussion? ↑
 [3081.10] 3C (-.57) That means, (0.16) I come in; now, because
 [3082.50] 3H (-.53) I'm ok;
 [3083.39] 3C (0.36) earlier; I wasn't, allowed, to speak.
 [3083.56] 3H (-1.11) Alright!! (0.48) yes yes yes.
 [3085.20] 3C (-.24) ya? (0.12) ya.
 [3086.03] 3T (0.1) ↓ hehehe ↓

C now takes charge of the meeting with some resistance from H. Her complaint that she had been kept out elicits from H *Alright!!* with a duration of 0.68 seconds, the double exclamation mark indicating a fall from very high voice. We could, if we so wished, adduce considerable evidence from the orthographic transcript to indicate that this was an ill-

tempered meeting, with friction between C and H. But this is not the view C expresses. On the contrary, she comments from the chair on the cordial nature of the meeting:

- [3965.59] 3C (-0.26) ɔya.ə (0.04) I found it. (0.22) I found.
 [3966.71] 3H (-0.42) ɪhahaha,
 [3967.62] 3C (-0.84) this, very, cordial. (0.46) You were both, very cordi[ə:l]. you know?
 [3970.20] 3H (-0.29) ![hə]!
 [3970.50] 3C (-0.13) when y[u:] said >things, like ok. shall we average it. (0.16) ya. (0.41) both, were very, agreeab[ə:l]. (0.09) you know? (0.49) s[o:] (0.11) this; was a very, friendl[r:], (0.09) moderation.

Although C cites averaging as evidence of cordiality, on at least one occasion this averaging is done after T has already revised her mark to bring it close to H's mark. This could give the impression of a head of department using her power to override her junior colleague's marks. But far from indicating that she is dissatisfied with the meeting in any way, T's responses to H's suggestions indicate that she found the meeting a useful learning experience.

Either C is misjudging a meeting she has herself chaired and T makes inappropriate responses, or else a mere record of the words spoken is in this case potentially misleading. So perhaps the true temper of the meeting is reflected in the prosody. In fact much of the prosody does not help either, because it – especially the prosody used by C and H – is highly predictable, and does little more than guide the listener to syntactic and information structure. In these circumstances, we have to draw the conclusion that the cordial nature of the meeting was almost certainly not reflected in the linguistically organised material at all.

How people feel about each other in general, and about such context-specific things as marking discrepancies, may have less to do with information transfer than with the quality of the interaction. In a meeting in which the linguistic information is potentially threatening, it is all the more important to avoid factual and personal misunderstandings. This takes us out of the linguistic comfort zone into *terra incognita*. Laughter no doubt played an important role, as did aspects of prosody and voice quality that currently remain beyond our ability to analyse. What we can be reasonably confident about is the role of linguistically unorganised particles. Interaction involving particles is continuous throughout the meeting, and it is consistently at worst neutral, and in most cases it is positive. This would have been a very different meeting if instead of making careful and considered responses to information provided by T (and generating hesitation phenomena in the process), H had produced fluent and rehearsed justifications of her own marks. Although the relationships among the participants might appear to be long established, and not being negotiated at this meeting, it is as well to recall the position of Baxter and Montgomery (1996), who argue that relationships are not in fact fixed, but are constantly re-negotiated in communication. In Arundale's (2010) terms, talkers constantly project connection and separation. Previous work (Zuraidah Mohd Don and Izadi 2011) has located these projections in the linguistically organised material, but in the present data, they are more realistically located in attitudinally positive linguistically unorganised particles.

CONCLUSION

This paper began with the question of the importance of prosody and linguistically unorganised particles in human interaction (see Chen & Wan 2016, for the role that prosody plays in intercultural communication). The case we have put forward for particles is that they

are far from constituting a homogeneous group, and can be sorted into emotive particles, artefacts of vocalisation, and those concerned with the sharing of knowledge between speaker and addressee. They can also be classified according to their relationship to speech, whether or not they have phonetic form, and if they have phonetic form, whether or not they are phonologically organised.

The case for prosody is that when it co-occurs with particles it operates in a different way than when it illustrates syntactic or information structure or plays a role in turn taking. It appears to operate as an independent component of particles, using salience to draw attention, or lack of salience to accompany activities which are not intended as interruptions.

An important consequence for the representation of conversational data is that it is essential to include an accurate notation of the prosody of particles. This is probably more important than recording the prosody of syntactically organised material. An important consequence for the interpretation of conversational data is that it is essential to include particles in the analysis, because as in the case of the data examined here, just looking at the words spoken can give a false impression of the nature of the interaction between participants.

ACKNOWLEDGEMENTS

This work was supported by the University of Malaya Research Grants (Project numbers: RP014A-13SBS and RG331-11 HNE). The authors would also like to thank Dr Mohana Nambiar for the significant contribution to the project.

ENDNOTES

¹ The fact that these particles are conventional, subject to well-formedness conditions, and need to be learned means that they must be part of some kind of communication system; but *pace* Enfield et al. (2013, p. 353), this does not necessarily make them part of language.

² This was an interesting finding, but not surprising. Different but functionally equivalent intonation systems are to be found in England itself, e.g. in Liverpool (Knowles, 1975).

³ Not to be confused with *raised voice*, which includes increased loudness, and implies anger and a combative manner.

⁴ and who'll

⁵ Television cooks, for example, praise themselves by uttering [*m*:]! when tasting their own food.

REFERENCES

- Aijmer, K. & Simon-Vandenberg, A.-M. (2003). The discourse particle *well* and its equivalents in Swedish and Dutch. *Linguistics*. Vol. 41, 1123-1161.
- Armstrong, L. E. & Ward, I. C. (1926). *A Handbook of English Intonation*. Cambridge: Heffer.
- Arundale, R. B. (2010). Constituting face in conversation: face, facework and interactional achievement. *Journal of Pragmatics*. Vol. 42, 2078-2105.
- Barth-Weingarten, D., Dehé, N. & Wichmann A. (2009). *Where Prosody Meets Pragmatics* (pp. 107-128). Bingley, UK.: Emerald Group Publishing
- Baxter, L. A. & Montgomery, B. M. (1996). *Relating: Dialogues and Dialectics*. New York: Guilford Press.
- Beach, W. A. (1993). Transitional regularities for 'casual' "Okay" usages. *Journal of Pragmatics*, 19, 325-352.
- Bolinger, D. (1972). Accent is predictable (if you're a mindreader). *Language*. Vol. 48, 633-644.
- Bolinger, D. (1986). *Intonation and its Parts: Melody in Spoken English*. Stanford, CA: Stanford University Press.
- Bolinger, D. (1989). *Intonation and its Uses: Melody in Grammar and Discourse*. London: Edward Arnold.
- Chen, H.C. & Wang, Q. (2016). The effects of Chinese learners' English acoustic-prosodic patterns on listeners' attitudinal judgement. *3L: The Southeast Asian Journal of English Language Studies*. Vol. 22(2), 91–108.
- Chomsky, N. & Halle, M. (1968). *Sound Patterns of English*.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. (1969). *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

- Enfield, N. J., Dingemanse, M., Baranova, J., Blythe, J., Brown, P., Dirksmeyer, T. & Torreira, F. (2013). *Huh? What? - a first survey in twenty-one languages*. In M. Hayashi, G. Raymond & J. Sidnell (Eds.), *Conversation Repair and Understanding* (pp. 343-380). Cambridge: Cambridge University Press.
- Falk, D. (2000). Hominid brain evolution and the origins of music. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music*. Cambridge, MA: MIT.
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*. Vol. 31, 931-952
- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics. 3: Speech acts* (pp. 41-58). New York: Academic Press.
- Guthrie, A. M. (1997). On the systematic deployment of *okay* and *mmhmm* in academic advising sessions. *Pragmatics*. Vol. 7(3), 397-415.
- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action. Studies in Conversation Analysis* (pp. 299-345). Cambridge: Cambridge University Press.
- Heritage, J. (1998). Oh-prefaced responses to inquiry. *Language in Society*. Vol. 27, 291-334.
- Heritage, J. (2012). Epistemics in action: action formation and territories of knowledge. *Research on Language and Social Interaction*. Vol. 45(1), 1-29.
- Heritage, J. (2013). Epistemics in conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 370-394). Chichester: Wiley-Blackwell.
- Heritage, J. & Clayman, S. (2010). *Talk in Action*. Sussex: Wiley-Blackwell.
- Jefferson, G. (2004). Glossary of transcript symbols with an Introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13-23). Philadelphia: John Benjamins.
- Kingdon, R. (1958). *The Groundwork of English Intonation*. London: Longman.
- Knowles, G. (1975). *Scouse: The urban dialect of Liverpool*. (PhD), University of Leeds, Leeds.
- Knowles, G., Williams, B. & Taylor, L. (Eds.). (1996). *A corpus of Formal British English Speech: The Lancaster / IBM Spoken English Corpus*. London: Longman.
- Lam, W.Y.P. (2009). What a difference prosody provides: The role of prosody in the study of discourse particles. In D. Barth-Weingarten, N. Dehé, & A. Wichmann (2009). *Where Prosody Meets Pragmatics* (pp. 107-128). Bingley, UK.: Emerald Group Publishing.
- Lieberman, A. (1996). *Speech: A Special Code*. Cambridge, MA: MIT Press.
- Local, J. (1996). Conversational phonetics: some aspects of news receipts in everyday talk. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation* (pp. 177-230). Cambridge: Cambridge University Press.
- Monetta, L., Cheang, H. S & Pell, M. D. (2013). Understanding speaker attitudes from prosody by adults with Parkinson's disease. *Journal of Neuropsychology*. Vol. 2(2), 415-30.
- O'Connor, J. D. & Arnold, G. F. (1973). *Intonation of Colloquial English*. London: Longman.
- Pierrehumbert, J. B. (1980). *The Phonology and Phonetics of English Intonation*. (PhD), MIT.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*. Vol. 1, 515-526.
- Reed, B. S. (2011). *Analysing Conversation: An Introduction to Prosody*. London: Palgrave.
- Sacks, H., Schegloff, E. & Jefferson, G. (1974) A simplest systematics for the organization of turn-taking in conversation. *Language*. Vol. 50, 696-735.
- Schank, R. & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structure*. Hillsdale, NJ: Lawrence Erlbaum.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Steele, J. (1775). *The Melody and Measure of Speech*. Menston: Scolar Press.
- Ting, S-H., Mahanita Mahadhir & Chang, S-L. (2010). Grammtical errors in spoken English of university students in oral communication course. *GEMA Online® Journal of Language Studies*. Vol. 10(10), 53-70.
- Tomasello, M. (2010). *Origins of Human Communication*. Cambridge, MA: MIT Press.
- Wells, R. S. (1945). The pitch phonemes of English. *Language*. Vol. 21, 27-39.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.
- Yngve, V. (1970). *On getting a word in edgewise*. Paper presented at the Sixth Regional Meeting of the Chicago Linguistic Society, 567-577.
- Zuraidah Mohd Don & Ahmad Izadi. (2011). Relational connection and separation in Iranian dissertation defences. *Journal of Pragmatics*. Vol. 43, 3782-3792.

APPENDIX 1

Transcription Symbols (adapted from Jefferson, 2004) including the finer phonetic details captured by way of symbols:

[]	Overlapping talk
(0.28)	Silence within turns or in talk
(.)	Discernable silence but less than 0.2 of second
[2876.78]	Time elapsed since the beginning of the sound file
.	A fall into the lowered range
,	Low rise
:	Higher rise (and often crescendo) into the middle range
!	A fall from high pitch
?	A high rise
?!	A high rise fall
.,	A low fall rise
↑	The beginning and end of high voice
!...!	The beginning and end of speech using an ingressive air stream
: (e.g. m:, a:)	Elongation of preceding sound
[h]	In breath
..	Breathy voice
~	Creaky voice
◦	Quiet voice
^N	Nasal voice
_L word _L	Laughter