

Extracting Meaning From Context: Modeling the Prosody of *Oh* in Mandarin Conversation

LI-CHIUNG YANG
Tunghai University, Taiwan

0. Introduction

The meaning and significance communicated between participants in normal spoken conversation is not only a function of the literal information transmitted, but is also dependent upon initial speaker knowledge states, level of interest and involvement in topics, and emotional response to the specific nature of the information exchanged. The co-existence and changing nature of these elements throughout the course of a conversation requires constant mutual signaling and coordination as uncertainties of fact and emotional response arise, are detected, and are resolved satisfactorily. Discourse markers play a pivotal role in this process, as they act to communicate the full spectrum of informational adequacy and personal emotional significance at key information exchange points in conversation. Because of the limited lexical content and short duration of discourse markers, prosody plays a prominent role in the successfully expressing the complex spectrum of meaning that is constantly at work in discourse.

The discourse marker *oh* is especially significant in the task of communicating meaning succinctly and accurately. Its universal presence in all languages and high frequency of use suggests that it has a basis in expression and communicating that is fundamental to the physiological and cognitive roots of human language. The meaning of *oh* arises in large part from its prosody because of its lexically limited nature and because of the close link between the cognitive, emotional, and physiological aspects of speech that it embodies. The discourse marker *oh* has been analyzed by Schiffrin (1987:100) as “a marker of information management” and as a signal of “shifts in speaker orientation (objective and subjective) to information.” In this paper, we focus on how the prosody of *oh* reflects cognitive and discourse phenomena of uncertainty and certainty and intensity of emotional response in expressing cognitive re-orientation. By examining the variability of form and function of prosody for *oh*, we hope in the current study to contribute to an understanding of how discourse context works together

with prosody to provide a more complete characterization of *oh* and its role in the communication of information and meaning.

1. Approach

In this study we present our results on prosody and discourse markers with a special focus on the prosody and discourse functions of the marker *oh*. We show how prosody contributes to the communication and interpretation of the multi-level meanings expressed in the marker *oh*, and suggest how the prosody of *oh* fits within an overall system of discourse. Our approach is to abstract both from the surface shapes and from knowledge of the specific emotional and cognitive processes, the specific discourse interactions and the topic structure, to form a theory linking the surface prosodic shapes to these other contextual or discourse elements. Our approach is productive in that we take an integrated approach of combining discourse analysis with a corpus-based approach, utilizing both acoustic and discourse data.

2. Methodology and Data

Our data consists of 2 sets of spontaneous conversations in Mandarin Chinese. There were 2 participants in each conversation, for a total of 6 hours of speech. Speech data were digitized using ESPS Waves+ software. These conversational data were analyzed and annotated for discourse relations, topic structure, discourse markers, and speaker turns. Instances of *oh* were extracted from the corpus and analyzed both acoustically and contextually. The acoustic measurements of pitch (f0), amplitude and duration of each instance of the discourse marker *oh* were correlated with the specific characteristics of the larger discourse context to derive characterizations of how the prosody of *oh* functions to communicate systematic variations in information state and cognitive expressiveness.

The results presented are based on a large number of instances of *oh* in different discourse situations, and derived from a very detailed analysis based on text, discourse context, and acoustic measures of prosody.

3. Discourse Markers and the Discourse Marker *Oh*

Discourse markers are words or phrases commonly used to signal the relationship between discourse units. They are words like *so*, *because*, *then*, *but*, and *now*, or feedback acknowledgements like *umhum* or *right*, and include particles like *oh*, *well*, and *hmm* (Schiffrin 1987, Heeman and Allen 1999). Discourse markers are significant elements of spoken language because they signal the flow of topic by acting as indicators between phrases and subtopics, and are critical for the ongoing interpretation of discourse. Discourse markers also serve as interactional devices guiding the exchange of information through the mutual coordination of knowledge states. Discourse markers tend to be highly expressive, because they occur at critical transition points of speaker intention and reaction.

The discourse marker *oh* is an immediate and primary marker of information state and emotion that occurs universally across languages. Its importance to

expressive speech and its occurrence as a critical focal point of information exchange make it a particularly significant subject for understanding the role and functions of human language. The universal appearance of *oh* in extremely diverse language families also suggests that study of this particle can provide potentially significant evidence on whether there is a systematic unified framework that can explain the relationship between prosodic form and the functional spectrum of *oh*. Such a framework can help us to obtain answers on the specific mechanisms of how prosody contributes to the communication and interpretation of the varied meanings that *oh* expresses.

4. Dimensions of Meaning: Surprise to Dawning Realization to Acknowledgement – the Marker *Oh*

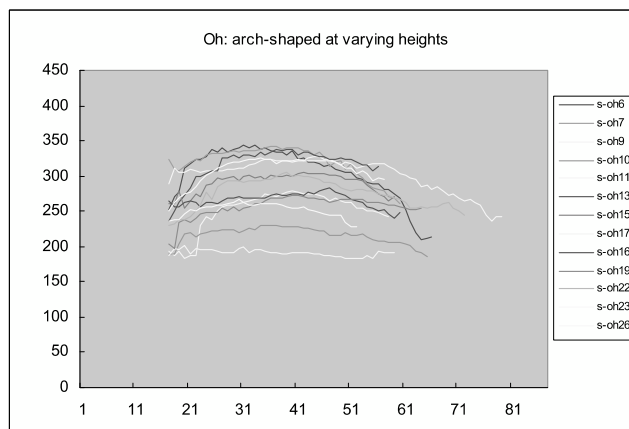
Oh is semantically unconstrained, so it is more transparent, iconic and expressive, and its meaning is derived from the context and the prosody. *Oh* can thus communicate a wide range of uncertainty-based states. Our analysis shows that one principle function of *oh* is as an expression of cognitive reorientation when encountering unexpected elements (cf. Heritage, 1998; Schiffrin, 1987). The cognitive reorientations can range from acknowledgement and acceptance to questioning, doubt, and disbelief, and are an important part of the interactive signals in discourse, as participants frequently use *oh* to signal alertness to new information or to indicate that the conversational exchange is proceeding smoothly. Moreover, upon encountering new information, participants often have both a cognitive reaction as well as an emotional response. The specific prosodic shape of *oh* is highly expressive in providing such types of information in an immediate, transparent way.

4.1. The Expression and Resolution of Uncertainty: Dawning Realization

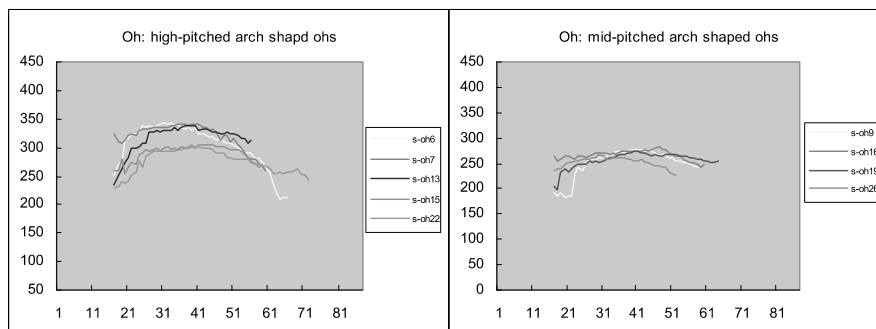
One of the most conspicuous uses of *oh* in discourse is as a marker of major cognitive readjustment after a period of uncertainty and miscomprehension, archetypically realized in the expression of “dawning realization.” In these cases, *oh* takes on a characteristic extended rise-fall shape, as exemplified in the instances of *oh* for speaker S in Figure 1, communicating different intensities of dawning realization. The extended rise-fall concave contours seen provide a picture of reorientation and development of cognitive state over time: rise is associated with surprise and uncertainty and fall with resolution and certainty, and both of these elements are at work in dawning realization. Contour variations of height, shape, and duration are critical elements in expressing the changing levels of uncertainty and certainty with respect to the information received, as well as the speaker involvement and level of emotional intensity.

The different pitch heights of Figure 1 reflect the intensity of initial surprise and status of the information with respect to the speaker’s knowledge state. High-pitch arched shapes reflect a high degree of initial uncertainty and subsequent high degree of resolution and acceptance.

Figure 1: Varying heights correlated with varying degrees of dawning realization



Figures 2-3: High-pitch and mid-pitch arch shaped *ohs* of Speaker S



- (1) W: wo qu le zhihou cai zhidao Kengding
haiyou yige jiushi -
s-oh6 S: *oh* haiyou jie xialai de.... Only after I went there then did I find out
there's another conference at Kenting –
Oh there's another one afterwards...
- (2) S: Computational de hen duo ma?
W: Yeah computational hen duo o
s-oh7 S: *oh* Lots of talks on computational?
Yeah, lots of talks on computational!
Oh

In example 1, speaker W is discussing a conference that she had attended and provides new information that clears up some confusion on this point by speaker S. This is clearly indicated in speaker S's response *oh haiyou jixialai de* "Oh there's another one afterwards"(s-oh6). In example 2(s-oh7), speaker S prompts

Modeling the Prosody of Oh in Mandarin Conversation

for more specific information relevant to her own understanding and interest of the conference, and speaker W's answer provides the precise information needed for clarification. The involvement of both speakers in the topic is high at this point, and the intensity level and need for clarifying information are correspondingly great. These factors cause the high pitch levels and long durations for *oh* here.

(3)

W:	wo yige gege haiyou wo meimei, wo nege gege shi xue yi de, ta shi shuyu xinzhang neike	One of my brothers and my younger sister – my brother studied medicine, he's a heart specialist
s-oh19 S:	<i>oh</i>	<i>Oh</i>
W:	na wo meimei shi shuyu gonggong weisheng de zheyangzi	and my sister studied public health that's the situation

A mid-level pitch range of *oh* often reflects acceptance and registering of information with a lesser degree of surprise, as shown in Figure 3. In example 3, speaker W is talking about her family and realizes the need to add more background information, thus she backtracks and provides more detailed information. Speaker S responds to the clarifying information on a topic that is of less direct personal interest and the mid-level pitch in s-oh19 provides a signal of comprehension and moderate interest in the subject.

The next example, example 4, shows instances of *oh* at the opposite extreme from high excitement and involvement. In these cases, a matter-of-fact acceptance of information that offers little challenge to the speaker's knowledge state causes the pattern of nearly flat pitch slopes seen in s-oh10 and s-oh17, which are at the lowest pitch levels in Figure 1. Emotions that are closely related to acceptance, such as sympathy and approval, also tend to be expressed in a low pitch level.

(4)

W:	wo - ta - wo tingshuo ta qu Princeton Shibushi	I heard she - I heard that she went to Princeton, isn't that so?
s-oh10 S:	<i>oh</i>	<i>Oh</i>
W:	qu nali jiaoshu a	She went there to teach
S:	ta qu naban jiaoshu a ?	She went there to teach?
s-oh11 W:	dui	Yeah / Oh (ah)
S:	<i>oh</i> (ah)	

In example 4, both speakers are discussing the whereabouts of a friend. S-oh10 occurs near where speaker W provides some information as a clarifying and not completely certain proposition. It seems that the information has not yet been fully grasped or resolved or it may be because that information is presumed known information, and this contributes to the low pitch level of s-oh10. However, in the subsequent utterance, there seems to be some conflicting information as speaker S follows with a repetition of speaker W's statement to prompt and verify the information. This is shown in the high level arch of s-oh11 as greater resolution occurs.

4.2. Degrees of Uncertainty and Doubt

In contrast to the arch-shaped contours of dawning realization, rising contours with little final fall often correlate with a high degree of uncertainty, doubt, and questioning. The cognitive uncertainty is frequently accompanied by a high emotional intensity because of the increased interest that a hearer has towards the subject. The varying pitch levels, contour shape, and sharpness of pitch rise express the specific intensity and character of the cognitive uncertainty at work.

The dramatic rising pitch forms for the particle *oh* in Figure 4 demonstrate these striking effects of uncertainty, interest, and emotional involvement on prosodic form. The sharp rise to a very high pitch level of s-oh1 is the direct result of incomprehension. In this example (example 5), speaker K had mentioned *nege wenquan* ‘that Bath’, but speaker S responds with an expanded pitch for *oh* that ranges from 212 to over 410 Hz, and ends at the highest pitch level of these cases:

- (5)
- | | | |
|-------|--|---|
| | K: zheshi lingwai yige difang nege wenquan | This is a different place. It's called Bath. |
| s-oh1 | S: <i>oh?</i> | <i>Oh?</i> |
| | K: zhe bushi Oxford zheshi Bath, B-A-T-H,
nege difang jiao Bath | This is not Oxford, it's Bath, B-A-T-H
That place is called Bath |
| s-oh2 | S: <i>oh!</i> hen youming, dui dui dui | <i>Oh!</i> Very famous. Right, right, right. |

In S-oh12, the high and sharp rise within a very short duration seen in Figure 4 quickly expresses the surprise of speaker S, which is accompanied with her repetition of the surprising fact that makes explicit the information that she had been unaware of, and this is followed by the immediate response of speaker W to explain this information:

- (6)
- | | | |
|--------|---------------------------------------|--------------------------------------|
| | S: oh ni qu xiaweiyi a oh | oh you went to Hawaii oh |
| | W: nege shihou wo meimei zai xiaweiyi | My sister was in Hawaii at that time |
| s-oh12 | S: <i>oh</i> ni meimei zai xiaweiyi | <i>Oh</i> your sister was in Hawaii |
| | W: ta huiqu le | She went back already |

In each instance of sharp pitch rise in Figure 4, there is a high level of uncertainty that arises both from the surprising nature of the new information as well as the high interest level of the respondent. The rising pitch shape of the lowest contour in Figure 4, S-oh27, also exhibits a very strong signal of alertness, curiosity, and interest, but does not rise as high as the other cases:

- (7)
- | | | |
|--------|--|---|
| | K: oh zheshi Yingguo zuei da de baihuo
gongsi | Oh this is the biggest department store in
England |
| s-oh27 | S: <i>oh?</i> ey, zenmeyang, Yingguo nege dian
zenmeyang? | <i>Oh</i> ey what was that like? How are the
shops in England? |

Figure 4: Instances of *oh* for Speaker S showing varying degrees of rise



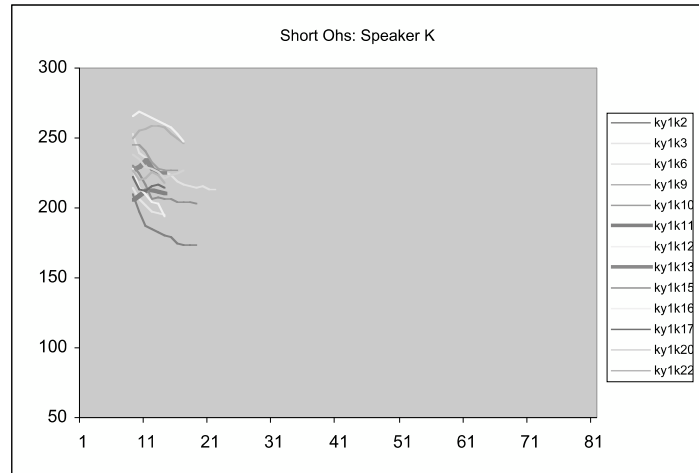
The contour of the rising pitch level in S-oh27 is itself adequate to express the questioning nature of the response. In this case, Speaker S’s alert *oh* response is elicited upon her hearing news about a very large department store, and the alertness and questioning is expressed explicitly in her follow up question, *ey, zenmayang?* “ey, what was that like?”, and the moderate pitch and gradual rise of s-oh27 expresses the speaker’s doubt and heightened interest. In both S-oh1 and S-oh27, the questioning intonation is strongly reinforced by the convex upwards pitch shape.

4.3. The Expression of Cognitive Certainty

We have seen examples of how the interjection *oh* acts as a marker of cognitive adjustment when new information is encountered in discourse, especially in cases of high uncertainty. In situations when the information encountered follows an expected pattern and does not cause significant surprise, the prosodic form of *oh* tends to be short and generally falling. While rising pitch is associated with uncertainty and greater cognitive adjustment, falling pitch contours are associated with certainty and less need for cognitive reorientation. In these types of situations, *oh* is usually used as a signal of acknowledgement of routine or logically expected information, or as a marker of comprehension of topic flow and information received. Figure 5 shows 13 instances of such short *oh*’s for Speaker K occurring at a range of different pitch levels:

In Figure 5, the falling contour instances of *oh* were predominately produced as responses to recollection of information or sudden realization of new information encountered, but with greatly varied degrees of intensity of interest towards the information. The higher the degree of intensity, the higher the associated pitch level. When *oh* is used as a marker of increased alertness and attention to new information, it is usually higher in pitch, with a short rise and fall, as is seen in several of the contours of Figure 5.

Figure 5: 13 Instances of Speaker K's short *ohs*



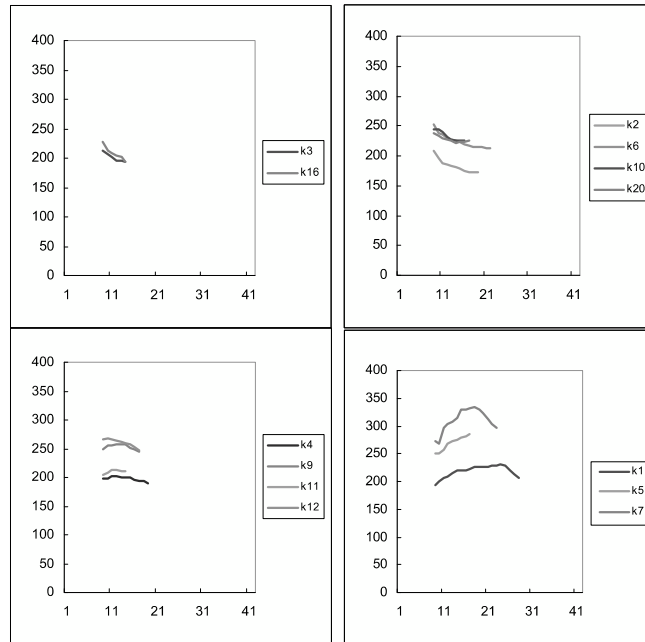
Figures 6-9 show the examples of Figure 5 by sub-type of pitch form, exemplifying expressions of differing but related cognitive states of *oh* for Speaker K. Although the instances of *oh* represented in these figures are all short in duration, the pitch patterns presented exhibit key characteristics that distinguish the expression of cognitive state that occurs in each case. In conversation, speakers frequently use *oh* to express simple acknowledgment, quick realization, or recollection of some related point or image, as in Figure 6 (upper left). In these cases, new information is encountered that is closely related to the ongoing topic, with little need for significant cognitive adjustment, and the prosody of *oh* is also correspondingly simple, with a very quick downward pitch that expresses the low degree of uncertainty of speaker K in those situations. Example 8 is a typical instance of dialogue that occurred:

- (8)
 k-oh16 K: *oh* zheshi lingwai yige difang nege wenchuan *Oh* this is another place – that's Bath

In Figure 7 (upper right), we see several further examples of sudden recollection by speaker K. Although the falling pitch *oh*'s displayed here are still quite short, the variations in pitch movement turn out to be very important. In these instances, the recollections occurring for the four instances of Figure 7 are all pleasant ones for the speaker, with little cognitive mismatch or effort. Because of this more pleasant context, the falling pitch shapes in Figure 7 have a more convex shape and slightly longer duration, aurally presenting a softer and more relaxed response.

Modeling the Prosody of Oh in Mandarin Conversation

Figures 6 – 9 (In left to right order): Breakout of speaker K's short falling *ohs*



(9)
 k-oh2 K: hen duo ne zhengge dixiashi *oh* A lot! The whole basement – *oh* we spent one
 kanle yige xiawu kan buwan tailei le afternoon looking, couldn't finish – too tired

The *oh*'s for k4 and k9 in Figure 8 (lower left) also have falling pitch and similar duration to those of Figure 7, but their contours have a more concave shape. The instances of *oh* in these cases were immediate responses to a perceived mismatch in knowledge states between the two speakers, as evidenced by the explicit correction that follows immediately in each case. The greater cognitive effort and recognition of mismatch that is present is expressed in the associated concave pitch shape.

(10)
 S: xian nadao degree meiyou? nian buo- Now has she gotten her degree yet? She's
 shi xuewei studying for her PhD
 K: umhum umhum
 S: nian wenxue fangmian de Studying literature
 k-oh4 K: *oh* buguo tamen nebian nian Ph.D. *Oh* but it must take you 10 years to get a
 yiding yao shinian Ph.D. there
 S: zhendea? Really?

Li-chiung Yang

- (11)
- | | | |
|-------|--|---|
| | S: ey ni na le meiyou? zhe zenme
zhiyou yizhang a? | Hey did you take it already? How come
there's only one here? |
| k-oh9 | K: <i>oh</i> yinwei wo – wo you yizhang jigui
ta le | <i>Oh</i> because I – I sent one to her already |

In contrast to the certainty of falling pitch, K's *oh*'s in Figure 9 (lower right) exemplify the effect of increased intensity, involvement and interest in a new topic. In both k5 and k7 in examples 12 and 13, *oh* has a dramatic initially rising shape, even though both are instances of recollection triggered by external and internal stimuli that are very relevant to the speaker:

- (12)
- | | | |
|-------|---|--|
| | K: hen nanchi (laugh) erqie erqie you hen
you hen gui | It tastes terrible! (laugh) also – also it was
really expensive |
| k-oh5 | <i>oh</i> nege shenme Deguo de nege McDon-
ald hen gui | <i>Oh</i> that – the McDonald's in Germany are
very expensive |

- (13)
- | | | |
|-------|--|---|
| | S: na haobuhao chi | Do they taste good? |
| | K: hen hao chi | Delicious! |
| | S: hen hao chi a | Really delicious? |
| | K: umhum keshi ta nege beer hall hen cao | Umhum but the beer halls were really noisy |
| k-oh7 | <i>oh</i> wo yao qu – wo yao huilai qian yitian
shi tamen Octoberfest | <i>Oh</i> the day before I went – I came back – it
was their Octoberfest |

The systematic variations in pitch shape that are seen in Figures 6-9 demonstrate that even within similarly short durations, contour differences are highly valuable in providing detailed information on a significant sub-spectrum of speaker meaning. Acknowledgment and attention response as seen here are frequently produced both as responses to information from the other speaker as well as by sudden retrieval of related stored information by the speaker herself.

4.4. The Effects of Co-occurring Emotions on Prosodic Form

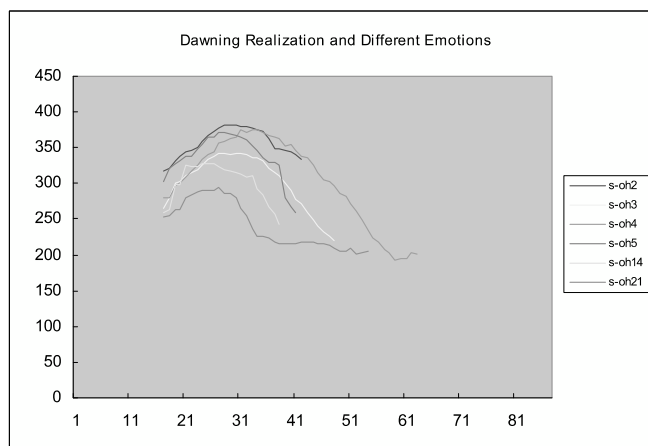
Responses to newly perceived information in discourse frequently contain both a cognitive reaction to the degree of unexpectedness of the new information as well as an emotional component that expresses the speaker's emotional evaluation of the specific information. Both components are expressed in prosody, and the co-occurrence of each virtually simultaneously gives pitch prosody a complexity and expressiveness that enriches the spectrum of meaning that discourse communicates so effectively. Figure 10 shows instances of surprise co-occurring with different emotions, each expressed within the duration of a single *oh* token.

Our analysis of the data shown in Figure 10 indicates that the high degree of uncertainty inherent in intense surprise causes a high rise in pitch, as in s-oh2, and s-oh5. The sharper and narrower arch shape of s-oh3 and s-oh4 reflect the presence of surprise with co-occurring emotions. As seen in example 14, speaker K is

Modeling the Prosody of Oh in Mandarin Conversation

talking about her impression of the Egyptian mummies at the museum, and speaker S responds with amazement and terror. The relatively shorter period of rise and extended low pitch level in the second half of *oh* in this case iconically expresses the fear and caution at work here. The narrow high arch of s-oh4 in example 15 reflects the speaker's amazement at the large number of colleges at Oxford, and the narrow arch that has a lengthy falling pitch to a low level in s-oh21 expresses both surprise and horror at a description of mummies in the museum.

Figure 10: Dawning realization in combination with different emotions-speaker S



- (14)
- | | | |
|-------|--|---|
| | K: ba nege ren na chulai ye he – heihei xiang
xiang nege rougan | They took the mummies out! It's black,
just like – like dried meat |
| s-oh3 | S: <i>oh-wa</i> | <i>Oh-wa!</i> |
| | K: rougan (laugh) D: rengan | Dried meat (laugh) D: Dried people |
| | K: dui, rengan. ta toufa dou hai you ye | Right, dried people. The hair is still there! |
- (15)
- | | | |
|-------|--|--|
| | S: nage-nega daxue zenmeyang, Oxford? | That - how was the university, Oxford? |
| | K: Oxford hen hen piaoliang | Oxford's really beautiful |
| | S: hen piaoliang ho | Really beautiful huh |
| | K: keshi ta limian you sanshi jige college | But it has more than 30 colleges |
| s-oh4 | S: <i>oh</i> hen da ma ho | <i>Oh</i> it's really big then! |
| | K: hen da | Really big |
- (16)
- | | | |
|--------|--|--|
| | K: ranhou haoduo dagai jiqian ge mummies
ranhou zhengzhengpai | Then there were many maybe thousands
of mummies and rows and rows of them |
| s-oh21 | S: <i>oh-wa</i> qeshi hen jinren | <i>Oh-wa!</i> that's really amazing |

Li-chiung Yang

6. Conclusion

In this paper, we have shown that that prosody plays an indispensable role in communicating the multi-dimensional contextual meanings carried by the discourse marker *oh*. We have shown that variations in overall pitch direction, duration, and specific shape correlate systematically with changes in the level of uncertainty and the nature of emotions that develop throughout natural discourse. The characteristic functions and prosody of *oh* presented in this study suggest that the marker *oh* presents a system of discourse contextual meaning in a crystalline and compact form, containing the essential elements of topic coherence, cognitive state, and emotional expressiveness that prosody brings to an entire discourse. We conclude that the ability of prosody to simultaneously signal both the cognitive degree of uncertainty and a specific emotional reaction to new information within the short time-scale of discourse markers such as *oh* is convincing evidence of the powerful expressive role of prosody in interactive discourse.

References

- Aijmer, Karen. 1987. Oh and ah in English conversation. In W. Meijs, ed., *Corpus Linguistics and Beyond*, 61-86. Amsterdam: Rodopi.
- Chafe, Wallace. 1992. Prosodic and functional units of language. In J. Edwards and M. Lampert, eds., *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, N.J.: Lawrence Erlbaum.
- Fischer, Kerstin & Hans Brandt-Pook. 1998. Automatically disambiguating discourse particles. *Proceedings of Coling/ACL '98 Workshop on Discourse Relations and Discourse Markers*, Montreal, Canada, 107-113.
- Heeman, Peter A. and James F. Allen. 1999. Speech repairs, intonational phrases and discourse markers: Modeling speakers' utterances in spoken dialog. *Computational Linguistics*, Vol. 25(4):1-45.
- Heritage, John. 1998. Oh-prefaced responses to inquiry. *Language in Society* 27:291-334.
- Schiffrin, Deborah. 1987. *Discourse Markers* [Studies in Interactional Sociolinguistics 5]. Cambridge: Cambridge University Press.
- Yang, Li-chiung. 2003. Visualizing spoken discourse: Prosodic form and discourse functions of interruptions. In J. van Kuppevelt and R. Smith, eds., *Current and New Directions in Discourse and Dialogue*. Kluwer Academic Publishers.

Li-chiung Yang
Department of Foreign Languages and Literature
Tunghai University
Taichung, Taiwan 407

Yang_lc@thu.edu.tw, lyang@sprynet.com