Tone and intonation in discourse management – How do speakers of Standard Vietnamese initiate a repair?

2 authors, including:

Martine Grice
University of Cologne
164 PUBLICATIONS   **2,906** CITATIONS

Some of the authors of this publication are also working on these related projects:

Brain Modulation and Speech Motor Control View project

Syllable structure and articulatory coordination View project

Tone and intonation in discourse management – How do speakers of Standard Vietnamese initiate a repair?

Kieu-Phuong Ha [*], Martine Grice

IfL - Phonetik, University of Cologne, Germany

Abstract

This paper investigates the intonation of other-initiation of repair in Standard Vietnamese and provides an analysis of its interaction with lexical tone at the word level. Results from both single-word and multi-word utterances show that speakers make predominant use of a rising pitch contour to mark other-initiation of repair, a contour that has been found in a large number of languages. This rise occurs consistently at the right edge of utterances and may be analysed as a phonetic reflex of a high boundary tone that can overlap, partially or completely, with the lexical tone of an utterance-final word. Further, the study demonstrates the application of talk-in-interaction to the study of the relation between lexical tone and intonation in tone languages.
© 2016 Elsevier B.V. All rights reserved.

Keywords: Intonation; Other-initiation of repair; Lexical tone; Talk-in-interaction; Vietnamese

1. Introduction

In conversation, when something that a speaker has said is misheard or when a speaker has problems understanding a previous utterance, an attempt is often made to elicit a repetition, clarification, elaboration or correction, referred to in Conversation Analysis as other-initiation of repair (henceforth OIR) (Schegloff et al., 1977). In many languages, repair is initiated by means of a distinct rising intonation (Enfield et al., 2013; Dingemanse et al., 2013). Our aim is to explore how repairs are initiated in Vietnamese, a language with a complex lexical tone system and extensive use of particles, factors that might be expected to leave little room for intonation to express communicative functions.

This paper investigates how speakers of Standard Vietnamese (also called Northern Vietnamese or Hanoi Vietnamese)[1] initiate a repair in conversation. Specifically, we are concerned with how speakers use intonation, and the extent to which this intonation interacts with the lexical tone of the words used for OIR. The study combines the methodology of Conversation Analysis, using naturally occurring data subjected to sequential analysis, with methods from Phonetics and Phonology, including supplementary elicitation methods, aiming at an in-depth analysis of the intonation used in the investigated context, especially in terms of its interaction with lexical tones.

---

* Corresponding author.
E-mail addresses: hak@uni-koeln.de (K.-P. Ha), martine.grice@uni-koeln.de (M. Grice).

[1] Northern Vietnamese is one of three main dialects in Vietnam, together with the Central and Southern Vietnamese dialects. Their tone systems differ not only in the number of lexical tones (six in the Northern dialect, five in the Southern and four/five in the Central variety), but also in the extent to which non-modal voice is used (Kirby, 2011; Brunelle, 2009a; Honda, 2006).

This study contributes to an increasing body of research on repair across languages (Hayashi et al., 2013; Enfield et al., 2013; Dingemanse et al., 2015), in particular on the prosody of this action in conversation (e.g. Couper-Kuhlen, 1992 for American English, Selting, 1996; Egbert et al., 2009 for German, Enfield et al., 2013 for 21 different languages, Dehé, 2015 for Icelandic). More importantly, it sheds light on the research area of talk-in-interaction as a sound foundation for intonation research in Vietnamese, an area that has so far received little attention among the research community of the language.

Below we will sketch the lexical tone system of Standard Vietnamese and provide an overview of how intonation has been described to date in Vietnamese in general and in this variety in particular (section 2) before addressing our research questions. After illustrating our data collection and methods (section 3), we will present the intonation found in OIRs, touching on its interplay with the lexical tones of the words under investigation (section 4). The question of how exactly the intonation of OIR interacts with lexical tones of the words will be the topic of section 5. Finally, we discuss the results and conclude in section 6.

2. Lexical tone and intonation in Standard Vietnamese

2.1. Lexical tone

Standard Vietnamese has six distinctive lexical tones, all of which occur on syllables that are either open or have a sonorant coda.[2] Some of these tones are characterised not only by pitch, but also by voice quality and duration (Alves, 1997; Nguyễn and Edmondson, 1998; Phạm, 2003; Michaud, 2004). This has led to a phonological analysis comprising both a pitch contour and a register for each tone, the latter involving laryngeal features, either glottalisation/creaky voice, or breathiness. Table 1 below provides an illustration of the six lexical tones on the syllable *la* [la:] along with their stylised contours, tone descriptions, traditional names and glosses.

In our analysis below, the target words in OIRs involve the lexical tones *ngang* (high-level), *ngã* (broken), *nặng* (low-glottalised), *hỏi* (low-falling)[3] and *huyền* (falling). Further details on the target words will be provided in sections 3.3 and 3.4.

2.2. Intonation

Earlier descriptive studies on Vietnamese intonation have identified a number of intonation patterns that are associated with functions in discourse. For Standard Vietnamese/Hanoi Vietnamese, Thompson (1965) proposes four *local* intonation contours occurring on the final syllable of utterances: (i) ''Decreasing'' intonation, which conveys that the speaker has not finished speaking and that he or she has been interrupted or has simply left something unsaid; (ii) ''Increasing'' intonation, which conveys ''exclamative statements and commands, and questions to which the answer is of relatively small interest to the speaker'' (ibid: 107–108); (iii) ''Fading'' intonation, ''including commands, statements of fact which the speaker assumes the hearer will accept as true, and questions to which the speaker assumes the answer'' (ibid: 108); and (iv) ''Sustaining'' intonation, used when the speaker is ''unable or for some other reason fails to predict the result of his speech'' (ibid: 108) or in ''doubtful statements, emotionally coloured and indecisive imperatives and exhortations, and informational questions'' (ibid: 109).

A survey of the Southern variety identified three basic intonation contours, marking different sentence types (Trần, 1967): The ''falling'' contour, which occurs in declarative statements, indicating that the speaker has completed his/her sentence; the ''increasing'' contour, which occurs mostly in interrogative and imperative exhortative sentences; and the ''sustaining'' contour, which occurs in sentences that are non-final. Crucially, Trần regards these intonation contours as *global* in nature (i.e. stretching over the whole utterance rather than being confined to the final syllable, as in the work of Thompson). Trần is also explicit about the fact that although intonation can modify the lexical tones phonetically, this modification does not imply a change in tonal category; tonal cues to the meaning of words are preserved (Trần, 1967: 23).

Following these earlier studies by Thompson (1965) and Trần (1967), an increasing number of investigations into the intonation of Vietnamese have employed experimental methods. For Standard Vietnamese, Đỗ et al. (1998) and Nguyễn and Boulakia (1999) identified a fast tempo, a global rising F0 and a local final rise in F0 as phonetic characteristics of

---

[2] Note that Vietnamese has also been analysed as having two additional tones, which are restricted to checked syllables, i.e. syllables ending with stop consonants (see e.g. Kirby, 2011). Since they are generally regarded by native speakers to be variants of the two tones *sắc* and *nặng*, and since the target words in the current study only involve syllables ending with sonorants, these two additional tones are not included in our representation of the tone system in this paper.

[3] The full form of this tone can be produced in isolation with a falling-rising contour. However, like tone 3 in Mandarin Chinese, it is predominantly produced as falling in phrase-medial position and can even take this form in isolation. This tone is often accompanied by breathiness (Phạm, 2003). Based on the most common acoustic properties of this tone, and following Kirby (2011), we categorise it as a low-falling tone.

Table 1

Stylised contours and tone descriptions of the six lexical tones in Standard Vietnamese (adapted from Kirby, 2011) based on the open syllable *la*. Diacritics on *la* are part of the orthography and are used to represent the tones. The high-level tone has no diacritic. Tones marked in grey are accompanied by non-modal voice quality (glottalisation for *ngã* and *nặng* or breathiness for *huyền*).

Stylised contour	Tone description	Traditional name	Example (orthographic)	Gloss
········	high-level	*ngang*	*la*	bawl
·····•••`	rising	*sắc*	*lá*	leaf
····?•·``	broken	*ngã*	*lã*	unclean
···· ?	low-glottalised	*nặng*	*lạ*	strange
·····``	low-falling[3]	*hỏi*	*lả*	exhausted
·········,	falling	*huyền*	*là*	to be

interrogatives, as opposed to the reverse patterns in declaratives. Vũ et al. (2006) and Nguyễn (2004) also identified a local final rise in interrogatives, even in the presence of a sentence final particle functioning as a question marker.

Recently, Brunelle et al. (2012), investigating the speech of 16 native speakers of Standard Vietnamese, found that intonation can be used to distinguish utterance types (interrogative vs. declarative) and affect (neutral vs. marked), with a tendency for higher pitch to be used in interrogative and marked utterances. However, their results revealed a considerable amount of inter-speaker variation. Given this variation, and the fact that postlexical/pragmatic functions of intonation are largely expressed by particles, reducing the need for intonational marking, they challenged the linguistic status of intonation in the language, especially with respect to its function in *grammar*. Nonetheless, concurrent investigations of short utterances from naturally occurring data have shown that lexical tones can be dramatically influenced by the intonation used in *interactional contexts* (Hạ and Grice, 2010; Hạ, 2010, 2012, 2013): Single-word backchannel utterances signalling attention or confirmation were consistently produced with a low or falling pitch, analysed as reflecting an intonational low boundary tone (L%). This intonational tone is able to override the lexical tone of a word. For example, the high toned word *vâng* 'yes' is produced with a falling pitch when used with the meaning 'yes, I'm listening' or 'yes, I agree'. These findings show that intonation plays an important role in Vietnamese discourse, regardless of the fact that the language is a tone language and that it makes extensive use of sentence particles. They also encourage further investigations into other interactional contexts in Vietnamese talk-in-interaction.

Based on the results from single-word backchannels (Hạ and Grice, 2010) and results from studies on question intonation in Vietnamese (Đỗ et al., 1998; Nguyễn and Boulakia, 1999; Nguyễn, 2004; Vũ et al., 2006; Brunelle et al., 2012) as well as studies investigating the intonation of repair across languages (Couper-Kuhlen, 1992; Selting, 1996; Egbert et al., 2009; Enfield et al., 2013; Dehé, 2015) our paper addresses the following research questions:

I. Are single-word OIRs in Vietnamese expressed by intonation (instead of through/in addition to the use of sentence particles) even if the words used for repair have their own lexical tones? If so, what kind of intonation contour(s) can be found and how do these contours affect the lexical tones?
II. If intonation on single-word OIRs is found, what happens in longer utterances, i.e. in multi-word utterances (e.g. 'He said what?' or 'Write down the situation of whom?')? What is the domain within which the intonation has an effect?

In the next section we provide an overview of our data collection and methodology used for analysis to address the questions above.

3. Corpus

3.1. Data collection

The speech materials in the current study are based on spontaneous speech and read speech, the latter serving as a control for the analysis of tone–intonation interaction. The spontaneous speech consists of a corpus of 18 telephone conversations with a total duration of approximately one hour. The telephone conversations were recorded by a total of 27 participants (12 men and 14 women, aged from 11 to 63, and one child aged 5) and were calls to friends, family members or acquaintances in informal and natural settings. Speakers recorded their own conversations after receiving instructions on how to use the technical equipment, which included a Digital Audio Tape Recorder or an Edirol Digital Recorder, and a microphone fixed to the landline telephone in speakerphone mode. The supplemental read speech, featuring target words produced in carrier sentences, was recorded by six of the participants who took part in the telephone recording sessions (four females and two males, aged from 11 to 60). Speakers were asked to read aloud only the target words that they used in their OIRs in telephone calls, repeating each word three times (see also section 5.1 for more information). This data allows for a direct comparison between the target words in spontaneous speech and in their citation form. Recordings of read speech were carried out by the first author. All recordings were made at a sampling rate of 16 kHz. For telephone conversations, signals were low-pass filtered at 4 kHz. These recordings were made with only one channel, i.e. the voice of the speaker at the other end of the telephone line was recorded in speakerphone mode. The overall acoustic quality of the acquired speech signal was sufficiently good for a reliable F0 contour analysis. Permission from participants to take part in the recordings was obtained orally (in the case of the child, permission was obtained through consent of the parents). In the analysis presented below, speakers are coded by gender (M for male vs. F for female) and initials, e.g. M_DA or F_KT.

3.2. Methodology

Our methodology consists of three steps: the identification of possible repair contexts, the analysis of pitch contours and the analysis of the interaction between intonation and lexical tone.

First, in order to identify instances of OIR in the telephone corpus and to categorise them according to source of problem in discourse, we adopted the methodology of Conversation Analysis (Sacks et al., 1974; Schegloff et al., 1977), by first transcribing the conversations orthographically and consequently building a collection of excerpts which exhibit potential OIR contexts. We then analysed the sequential organisation of these excerpts, paying careful attention to how an utterance is responded to in the next turn. When labelling instances of OIR, we apply the repair structure annotation suggested by Enfield et al. (2013: 346) which consists of labelling three turns:

1. T−1: the trouble source
2. T0: the turn of other-initiation of repair
3. T+1: the turn of repair

This turn structure is made explicit in all examples/excerpts below.

Second, we analysed the pitch contours (and fundamental frequency traces) in the investigated OIRs and identified the lexical tone on each word occurring in these contexts (and in the case of multi-word OIRs, the final word), in order to perform a subsequent analysis of the interaction between intonation and lexical tones. The F0 contours were extracted in Praat (Boersma and Weenink, 2015), corrected using a customised version of mausmooth (Cangemi, 2015), and plotted in Praat and R (R Core Team, 2015).

Third, we analysed the interaction between lexical tone and intonation. As our target utterances were taken from natural speech, instances were necessarily limited in number and scope, and, moreover, were produced by different speakers. The analysis of tone–intonation interaction thus has a more qualitative than quantitative nature. The following analytical methods were used to explore the nature of tone–intonation interaction:

 i. Comparison between telephone data and elicited read speech, to identify the pitch contour expressing the context of OIR.
 ii. Comparison between utterances from telephone data only: target utterances vs. similar utterances used in a different discourse context, to see how they differ.
 iii. Analysis of deviant cases, to see what happens when the expected pitch contour is not used.

Method (i) is used to analyse tone–intonation interaction in single-word OIRs. We identified words used in single-word OIRs and supplemented the data with read sentences from six of the speakers who took part in the recordings of telephone conversations (see section 5.1). We then compared the lexical tone contours of the words used in OIRs with the

contours of the same words produced in citation form in the carrier sentence *Tôi dạy từ X cho bé* "I teach the word X to the child" (X = target word). The latter served as a baseline for the comparison. The carrier sentence was used to avoid word list intonation and utterance-final falling intonation.

Method (ii) is used to analyse tone–intonation interaction in multi-word OIRs. We compared the intonation patterns on final syllables in target utterances with those produced on final syllables in utterances in other discourse contexts. All data used in the comparison stem from the same speakers, have the same lexical tone, and are identical or similar in their segmental makeup. This analysis aims to provide supporting evidence for the intonation found in OIRs.

Method (iii) analyses so-called 'deviant cases' from the corpus, cases in which a repair is other-initiated with a different intonation contour from the identified/common pattern. Analysis of these cases demonstrates the relevance of the identified intonation pattern, showing that interlocutors respond differently when the common intonation pattern is not used (Ten Have, 2007). Methods (ii) and (iii) provide evidence for the form-function relation.

3.3. Example of single-word other-initiation of repair

We identified 40 single-word OIRs in the corpus. Extract (1) illustrates an OIR sequence due to a mishearing, a problem which can be identified based on what happens next in the conversation, usually an unproblematic repetition of the previous turn. A Vietnamese transcription is provided in italics, followed by a word-for-word gloss and a sentence translation in single quotes; see also Appendix A for abbreviations and conventions used in transcriptions and Appendix B for a list of particles used in the transcripts.[4] Mishearings constitute the majority of the target utterances in the corpus. The repair sequence in (1) consists of an overlap in speech (marked by square brackets) indicating the trouble source, T−1 (line 2), a next turn other-initiation of repair, *gì*, T0 (line 3, marked in bold), and T+1 (line 4), the turn in which the mishearing problem is resolved.

(1) Entrance exam ((M_HN and M_BN are friends, both aged 19))

```
     1   M_BN    tôi chả biết    [nó thi
                 I   NEG know  she take exam
                 'I don't know about her exam.'
     2   M_HN            [nó đô    trường    gì?                              ⌉
                         she pass university what                            │ T-1
                         'In which university did she pass the exam?'        ⌋
                 (0.9s)
 ->  3   M_BN    gì?                                                         ⌉ T0
                 'What?'                                                     ⌋
     4   M_HN    nó  đỗ  trường           gì?                                ⌉
                 she pass university      what                               │ T+1
                 'In which university did she pass the exam?'               ⌋
     5   M_BN    tôi cũng chả biết   nó  đỗ   trường gì    ý
                 I   also  NEG know  she pass univ.  what PART
                 'I don't know in which university she passed the exam either.'
```

Table 2 provides an overview of the tonal properties of the single-word tokens investigated in the analysis, along with their frequency in the corpus, their word classes and glosses. These tokens involved six lexical items and one non-lexical item: *ai* (who), *ơi* (vocative particle), *sao* (why/how), *gì* (what), *hả* (yes/no question word), *dạ* (yes – polite form), and *mm* (bilabial voiced nasal). The non-lexical item *mm* does not have a lexical tone and could be thus seen as an independent indicator of whether we are dealing with a particular intonation pattern that is used to initiate a repair.

3.4. Example of multi-word other-initiation of repair

We found 50 multi-word OIRs in the corpus. Twenty-seven of these were chosen for the F0 contour analysis and the investigation of the interaction between lexical tone and repair intonation.[5] Target utterances were produced to convey

---

[4] In cases where the word-for-word gloss is identical to the sentence translation, only the latter is provided.

[5] 23 tokens involve the final question particles *à* and *á*, transcribed in the orthography with a lexical falling tone and a lexical rising tone, respectively. It is unclear in Vietnamese whether the pitch contour produced on these particles are truly lexical tones, possibly resulting from grammaticalization of the intonation (as e.g. in Hong Kong Cantonese, Cheung, 1986; Wu, 2008) or whether the syllable is toneless, pitch contours being entirely due to intonation (as e.g. in Mandarin Chinese, Wu, 2006, or in Thai, Pittayaporn, 2007). Therefore, we exclude these 23 tokens from the analysis.

Table 2
Tonal properties of single-word OIRs, along with their frequencies, word classes and glosses.

Single-word OIR	Tone in citation form	Frequency	Word class	Gloss
ai	high-level	3	question word	who
ơi	high-level	2	particle	vocative
sao	high-level	1	question word	why/how
gì	falling	5	question word	what
hả	low-falling	23	question word	yes/no ques.
dạ	low-glottalised	5	particle	yes (polite)
mm	none	1	non-lexical	none
Total		**40**		

problems in understanding or mishearing due to previous overlapping talk, to check for understanding, to clarify/to request clarification of a co-referent or to request for elaboration (see the list of individual utterances according to problem source in Appendix C). Extract (2) presents an example of a multi-word OIR utterance. This excerpt stems from a conversation between two school friends, speakers M_DA and M_BDA; the former was reading aloud the homework for the literature class which the latter had not attended:

(2) Homework ((M_DA and M_BDA are school friends, aged 11))

```
  1  M_DA   ý nghĩa của các  chi tiết ( ) tượng (-)  kì ảo            ⌐
             meaning of  PL. detail   (ima)gination miraculous        |  T-1
             'Meanings of the details of miraculous (ima)gination.    |
             (3.1s)((M_BDA writes when talking))                      ⌐
-> 2  M_BDA  ý nghĩa c[ủa các   gì?                                   ⌐
             Meaning of  PL. what                                     |  T0
             'Meaning of what?'                                       ⌐
  3  M_DA              [của các                                       ⌐
                         of  PL.                                      |
                       'Of the'                                       |
             (0.5s)                                                   |
  4          ah      của ý nghĩa   của các chi tiết tưởng tượng kì ảo |  T+1
             Hesitation of   meaning of PL. detail   imagination miraculous |
             'Ah, no, meanings of the details of miraculous imagination.' |
             (1.8s)                                                   |
  5          đọc  thiếu                                               ⌐
             read lack
             'I didn't read it properly.'
```

The trouble source turn, T−1 (line 1), involves unintelligible material (indicated by empty brackets) read aloud by speaker M_DA to speaker M_BDA. This leads to the OIR in T0 (line 2) with the latter repeating a part of the turn by M_DA in line 1 and using the question word *gì* 'what' to indicate that he has misunderstood. In lines 3–5, T+1, speaker M_DA responds to the OIR in T0 by formulating the information more clearly, 'meaning of the details of miraculous imagination', and admitting that in his previous turn he had not articulated this clearly.

As it will become clear later in the analysis, the final syllable can be identified in multi-word OIRs as the domain within which intonation can affect the phonetic realisation of the lexical tone. Thus, in what follows, we concentrate on the tonal characteristics of the final syllable/word in multi-word utterances (see also Appendix C for complete target utterances). They involve not only sentence particles and functions words, as in the single-word utterances illustrated above, but, more importantly, also lexical words, whose tonal properties are reportedly more resistant cross-linguistically to the effects of intonation than those of particles or function words (cf. Selkirk, 1996). Table 3 provides an overview of the tonal characteristics and gloss of the final syllables occurring in multi-word OIRs and their frequencies in the corpus. These syllables involve four lexical tones: high-level, broken, falling and low-falling.

Table 3
Tonal properties of the final syllables in multi-word OIRs, along with their citation forms, frequencies, word classes and glosses.

Final syllable in multi-word OIR	Tone in citation form	Frequency	Word class	Gloss
ai	high-level	1	question word	who
cơ	high-level	5	particle	expression of intimacy or familiarity bet. speakers
sao	high-level	1	question word	why/how
nữa	broken	5	particle	else
gì	falling	10	question word	what
dòng	falling	2	noun	book line
nhà	falling	1	noun	home
hả	low-falling	2	question word	yes/no ques.
Total		**27**		

4. Intonation of other-initiations of repair in Standard Vietnamese

4.1. Single-word utterances

The pitch contours found in single-word OIRs are presented in Fig. 1, grouped according to word and lexical tone (high or low/falling).[6] Across the board, OIRs predominantly have a rising pitch contour. However, the details of this rising

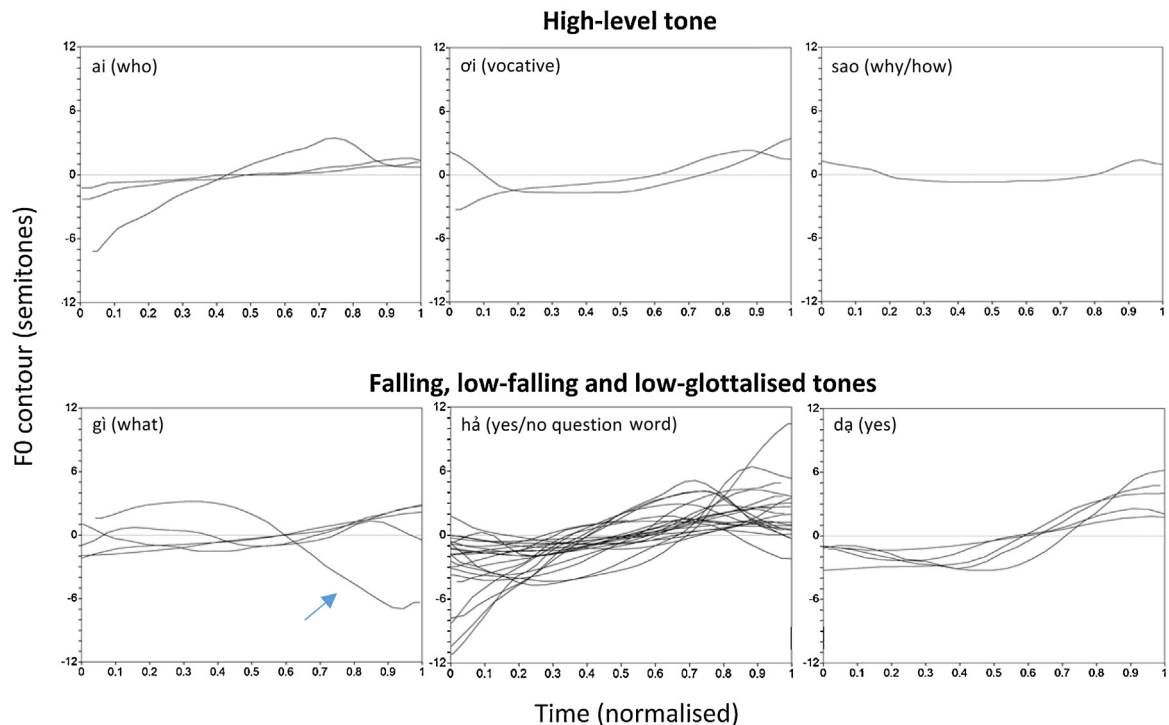Fig. 1. Normalised fundamental frequency contours of 39 single-word other-initiations of repair in the corpus for each investigated word: *ai* (n = 3), *ơi* (n = 2), *sao* (n = 1), all lexically high-level; *gì*, lexically falling (n = 5), *hả*, lexically low-falling (n = 23) and *dạ*, lexically low-glottalised (n = 5). The falling contour marked with an arrow will be returned to later and analysed as a deviant case.

---

[6] Our main concern here is the direction of the pitch movements in relation to those of the citation form of the lexical tones. As our data were produced by male and female speakers as well as a five-year-old child and these data involve different lexical tones, we plot the investigated pitch contours, following Nolan (2003), in semitones in relation to the mean of each contour separately, enabling the comparison between the contours across gender and lexical tones. This scale has been seen as "a well-motivated choice" for plotting intonation contours in terms of melodic equivalence (Nolan, 2003).
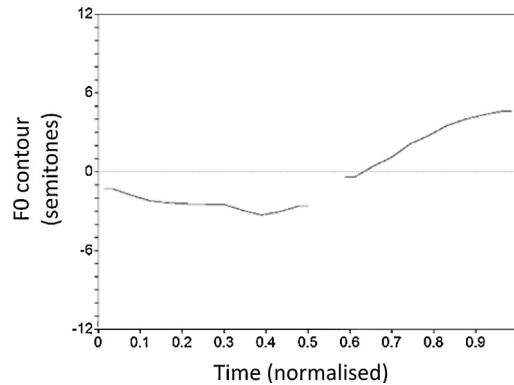
Fig. 2. The rising pitch contour produced on the word *dạ* ('yes' – polite) in an other-initiation of repair by speaker F_BB. *Dạ* is lexically low-glottalised (interruption in F0 contour indicates glottalisation).

contour depend somewhat on the identity of the lexical tone of the words involved. For the words with a high-level tone (first row), the pitch is either slightly rising or remains level. This is possibly due to a ceiling effect as the lexical tone of the words is already high. Interestingly, the low toned words (second row) are produced with a rising pitch contour, rather than their lexically determined fall. There is only one exception (on the word *gì*, lexically falling, marked with an arrow in the figure), which retains its falling contour. We return to this case in section 4.2. With respect to the contours of the question word *hả*, some appear to have a fall in the last 30% of the signal. However, there is a clear rise over at least the first 70% of the syllable, and the low-falling contour of the lexical tone is missing, arguably expressing OIRs. The final fall here might be caused by a number of factors involved in conversation, such as the attitudes of the speakers, the relationship between them (cf. Mạc et al., 2010), or possibly the sequential context in which the OIR occurs.

The most clear case of intonation overriding the lexical tone is the OIR used with the word *dạ* 'yes' (Fig. 1, last box, second row) which has a lexical low-glottalised tone, i.e. a fall followed by a glottal stop. In four out of five examples from the corpus, there is no glottalisation, i.e. only one OIR retains this feature. In Fig. 1 we plotted all the contours as smoothed trajectories interpolating between values. Fig. 2 shows the one example which retained glottalisation without smoothing in order to have a closer look at the interaction between the pitch contour and glottalisation. In general, for utterances with *dạ* 'yes', different pitch contours can lead to different meanings in discourse: if the utterance is produced with a low pitch ending with a glottal stop (corresponding to its lexical tone low-glottalised *nặng*), it conveys acknowledgement 'Yes.' (see also Table 2), and if it is produced with a rising pitch, with or without a glottal stop, it initiates a repair, 'sorry?' or 'pardon?', asking for repetition or elaboration. Interestingly, as shown in Fig. 2, the glottal stop or glottalisation was not final, but rather around the middle of the word. The pitch contour following it is rising.[7]

The non-lexical item *mm* also has a rising pitch as illustrated in Fig. 3. Since the item *mm* does not have a lexical tone, the pitch contour found on this item may be taken as evidence that a rising intonation is used to other-initiate a repair (especially since *mm* can also be used as a backchannel with a falling pitch, Hạ, 2010; Hạ and Grice, 2010).[8]

In sum, nearly all investigated utterances have a final high pitch or a final rise in the F0 contour, regardless of the lexical tone of the syllable. Only one case was found with a falling pitch. We discuss this case in the next section.

4.2. A deviant case

In extract (3) below we present the example of the OIR *gì*, lexically falling, which is analysed as a deviant case (see Fig. 1, contour marked with an arrow), a case that does not have the rising contour identified for OIRs. The analysis of this case can demonstrate the relevance of the identified rising pitch pattern in OIRs, in that different reactions of interlocutors can be observed when this pattern is not used (Ten Have, 2007).

The sequence of this conversation is between two colleagues, M_AK and F_CT (the latter older than the former). The extract begins with M_AK asking for feedback on how much should be withdrawn from the account he manages for paying

---

[7] This result corresponds to the results found in Nguyễn et al. (2013) that "glottalization is phased earlier for surprise [corresponding to OIR] than for declaration".

[8] Dingemanse et al. (2013) report that speakers produce the bilabial voiced nasal *mm* when they are close to each other, in terms of their physical distance. The reason why the amount of these items is small in our corpus seems to be related to the fact that speakers communicated via telephone and that their physical distance was not as close as when speakers communicate face-to-face.
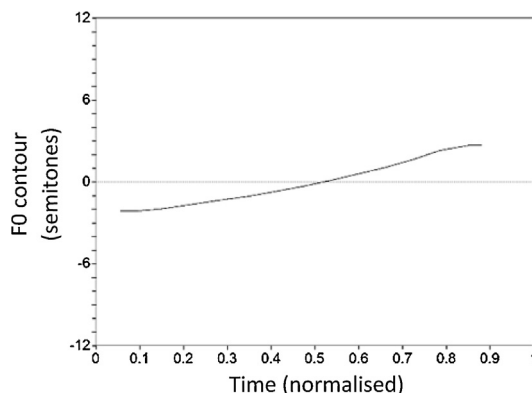
Fig. 3. The rising pitch contour produced on the non-lexical item *mm* in an other-initiation of repair by speaker M_BDA.

salaries to employees. The trouble source turn (T−1) is displayed in line 7, where speaker M_AK asks speaker F_CT about the amount of money she needs in order to pay the employees. The latter appears to be surprised because she does not expect the amount to be so high (more than 20 million VND) and initiates a repair (T0) by saying "What!" with a falling intonation (line 8). Her request for elaboration in this turn indicates her disbelief/surprise which can be observed from the *response of the recipient*, her colleague M_AK, to her OIR in line 8. M_AK not only specifies the amount of money (25 million) but he also uses the final particle *mà* emphasising that this information should have already been known to F_CT.

The disbelief/surprise of speaker F_CT is clear from her pauses during M_AK's elaboration (T+1), e.g. after line 9, line 10 and line 11. The last pause lasts nearly 3 s before F_CT gives the instruction to withdraw the money to pay the employees (line 12).

(3) Colleague ((M_AK, F_CT are colleagues, aged between 30-45))

```
  1  M_AK    Nhưng mà có rút        nhiều không.  rút        hết không?
             But      CÓ withdraw much KHÔNG withdraw all KHÔNG
             'But should I withdraw a large amount, should I withdraw all of it?'
             (1.1s)
  2  F_CT    Ờ   mày rút       hết cả cái    tiền    ấy     đi    chứ!
             yes you withdraw all   CLASS money 3. Ps. PART PART
             'Yes, withdraw all of it!'
  3  M_AK    [rút        hết
             withdraw all
             'All?'
  4  F_CT    [tiền công phải rút        hết để        trả nó   chúng nó chứ.
             Salary     must withdraw all in order to pay him them     PART
             'We need to withdraw all the money to pay them.'
  5  M_AK    rút        hết cả tiền lương hả?
             Withdraw all     salary        yes/no ques.
             'Withdraw all salaries?'
             (1.1s)
  6  F_CT    không. tiền lương thì     để     đấy cũng được.
             No      salary     PART leave there also   OK
             'No, you can leave the money for salaries as they are.'
  7 M_AK     nhưng mà tiền     đấy cô ( ) vẫn nợ  à (.)       hai    mấy     triệu đấy.
             But       money that aunt   still owe yes/no ques. twenty something million  PART
             'But you haven't paid that money (the salaries for workers), have you? Over 20 million
             dong?!'
             (1.3s)
-> 8 F_CT    GÌ!
             What!
```

T-1

T0

9 M_AK *hai nhăm triệu mà!*
      Twenty five million PART
      '25 million dong, you know?!'
      (0.9s)

10   *nhưng cô có dùng tiền của cô không rút   ( )*
      but aunt CÓ use money of aunt KHÔNG withdraw ( )
      'But do you use your own money?'
      (1s)

11   *sẽ rút   khoảng mấy triệu cô bảo cháu cô tính   đi!*
      will withdraw about how many million aunt tell me aunt calculate PART
      'If we withdraw millions, you tell me how much, just calculate!'
      (2.8s)

12 F_CT *Mày cứ rút   cho cô đi!*
      You still withdraw for aunt PART
      'Just withdraw some for me!'

13 M_AK *[rút   tất á?*
      Withdraw all yes/no ques.
      'All?'

14 F_CT *[rút   hai mươi triệu đi!*
      Withdraw twenty million PART
      'Withdraw 20 million!'

15 M_AK *[rút   tất á?*
      Withdraw all yes/no ques.
      'Withdraw all of it?'

16 F_CT *[để   trả công cho nó   chứ.*
      in order to pay work for him/her PART
      'We need to pay them.'

17 M_AK *Cháu tưởng cô dùng tiền của cô cũ trả rồi.*
      I think aunt use money of aunt former pay PART
      'I thought you had used your own money to pay already.'
      (--)

18 F_CT *Ừ thì cô trả nhưng bây giờ mày rút   mày cứ rút hai mươi triệu   về*
      Yes that aunt pay but now you withdraw you still withdraw twenty million back
      cho cô!
      for aunt
      'Yes, I need to pay, but now you just withdraw 20 million for me!'

19 M_AK *ah   khoảng hai mươi triệu thôi chứ gì?*
      Hesitation about twenty million only PART
      'Hmm, only about 20 million?'

20 F_CT *Ừ.*
      'Yes.'

21 M_AK *có   hai nhăm triệu rút   hai mươi triệu nhá?*
      Available twenty five million withdraw twenty million PART
      '20 million out of 25 million, correct?'

22 F_CT *Ừ.*
      'Yes.'

23 M_AK *còn đâu thì   để lại   tiền tiền lương nhá?*
      Rest where FOC PART leave back money salary   PART
      'The rest we will pay for their salaries, yes?'

24 F_CT *Ừ.*
      'Yes.'

25 M_AK *rồi OK.*
      'Yes, ok.'

(T+1 bracket spanning lines 9–11)

This example provides evidence for intonation functioning as a contextualization cue in interaction (Gumperz, 1992). As seen in section 4.1, a rising contour is used to initiate a repair due to general problems such as mishearing, whereas the lack of a rise on exactly the same word is used to indicate disbelief. The pitch contour on *gì* (line 8) is falling, corresponding to the falling pitch contour of the lexical tone *huyền* of the word. This example shows that, in Standard Vietnamese, speakers can use a rising pitch or the lack of it to express different meanings in discourse.[9] Further work

---

[9] It is unlikely that this is an isolated exception, as this is also true for multi-word repair initiations dealt with in section 4.3 below.

needs to look at more data to ascertain whether this result corresponds to those found for English or German *What?* or *Was?*, respectively, which commonly express mishearing when produced with a rising intonation, while the falling intonation found on the same token *What!* or *Was!* expresses surprise or disbelief (Schegloff, 1997; Egbert et al., 2009).

4.3. Multi-word other-initiation of repair

Overall, we also found a rising pitch contour at the end of 25 multi-word utterances and the lack of a rise at the end of two utterances (the latter are analysed as deviant cases). The rising pitch appears to affect only the final syllable in a phrase. Below we first present individual examples of this pitch pattern before giving an overview of the pitch patterns measured in all investigated tokens. Fig. 4 presents four utterances as illustrative examples, the final syllable of each carrying a different lexical tone: (a) high-level, (b) falling, (c) low-falling, and (d) broken.

In (a) the lexical tones in the non-final syllables are apparent: *tô* (lexically high-level), *đậm* (lexically low-glottalised), *gì* (lexically falling), whereas there is a clear rise on the final syllable *cơ* (although it is lexically high-level), especially in the second half of this syllable. Characteristics of the lexical tone present in the first half of the syllable (level pitch) may possibly facilitate recognition of this tone, although progressive assimilation in Vietnamese means that the initial portions of syllables play a very minor role in tone recognition (Brunelle, 2009b).

In (b), the last syllable of the utterance has a lexical falling tone. The non-final words in this utterance contain the following lexical tones: *số* 'number' rising, *di động* 'mobile' high-level and low-glottalised, *hay* 'or' high-level, *số* 'number'
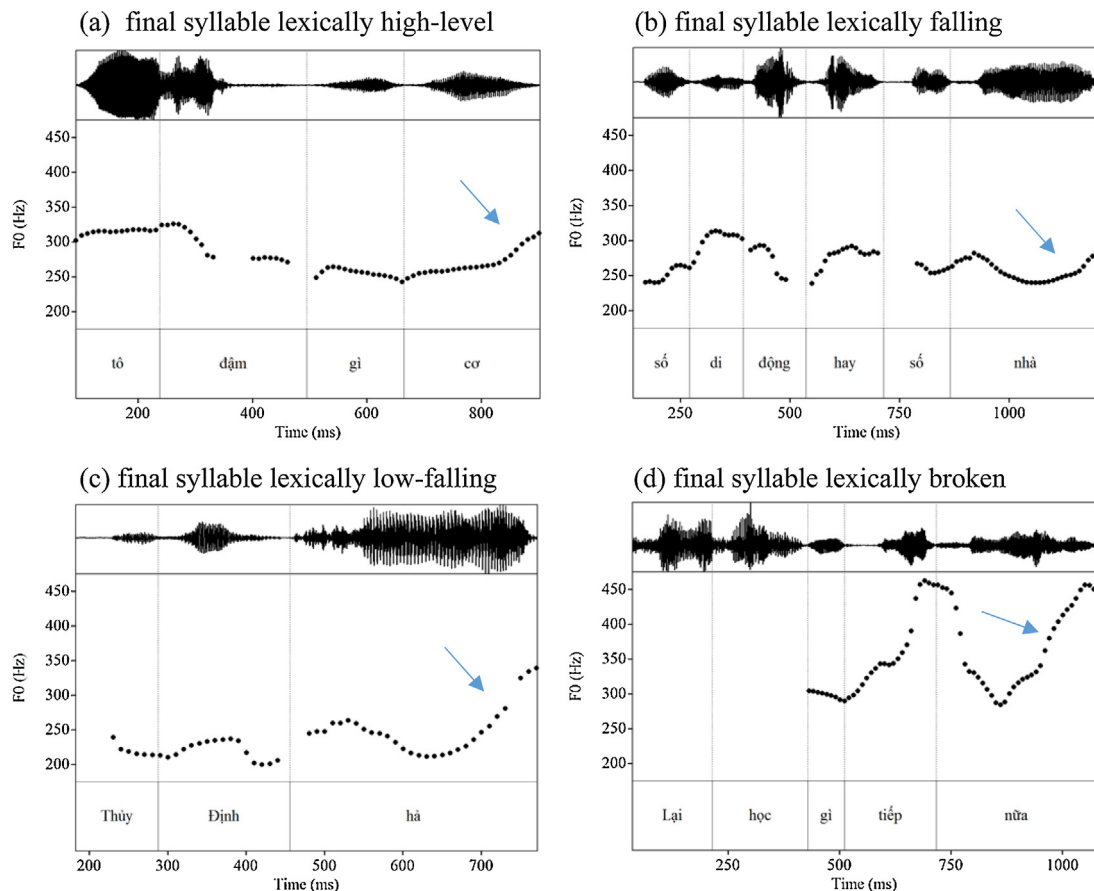

Fig. 4. Fundamental frequency contours produced on multi-word utterances, the final syllables of which have different lexical tones and belong to different word classes: (a) *cơ*, high-level, particle, (b) *nhà*, falling, noun, (c) *hả*, low-falling, question word, (d) *nữa*, broken, particle.[10]

---

[10] The beginning *Lại học* of the whole utterance in (d) *Lại học gì tiếp nữa?* 'What [does she want] to study further?' was produced in overlap with the turn of the other speaker, therefore the fundamental frequency contour of this part is not plotted here.

rising, and *nhà* 'home' falling. In this example too, the non-final words all retain their tonal characteristics, it is only the final word that appears to be modified. In this last syllable, the lexically determined fall is followed by a rise in the second half of the syllable.

In (c), the lexical tone of the word *hả* is low-falling. The intonational rise initiating a repair follows the (low-)falling contour of the lexical tone of this word, beginning halfway through the syllable. Again, the lexical tones of two preceding non-final words retain their lexical tone characteristics: *Thủy* 'person name' (low-falling), *Định* 'person name' (low-glottalised). The rising contour on the first part of the syllable *định* can be accounted for by progressive coarticulation from the tone in the preceding syllable.

In (d), the lexical tone of the final syllable is broken (rising with glottalisation), and in its canonical form (see Table 1) the glottalisation occurs halfway through the syllable. However, here, there is no evidence of glottalisation; instead there is an uncharacteristically high rise at the end of the contour.

Fig. 5 provides an overview of the pitch contours produced on the final syllables in 25 investigated multi-word OIRs, grouped by their lexical tones (see also Appendix C). As expected, the contours on the broken tone rise, as shown in (i). Interestingly, the contours on high-level tones also rise, as illustrated in (ii), although in some cases there may still be a ceiling effect preventing a further rise. For the falling tone (iii), most contours do not fall evenly throughout the syllable. Instead there is a rise predominantly during the second part of the syllable, an effect that is unsurprising if the rise is due to a phrase-final intonational edge tone. Only two cases have a falling contour, marked with an arrow. We will show below that these are deviant cases (akin to the single word deviant case discussed in section 4.2 above). For the low-falling tone in (iv), the contours also rise and the rise begins halfway through the syllable.

In sum, the rising intonation used in OIRs is found not only in single-word utterances, but also in multi-word utterances (except for two cases analysed in the next section). This intonation appears to be stable across speakers in this communicative context. However, there may be a ceiling effect in the case of the lexical high-level tone together with the OIR pitch contour, underpinning the fact that the manifestation of the rising intonation for OIRs is tone dependent.
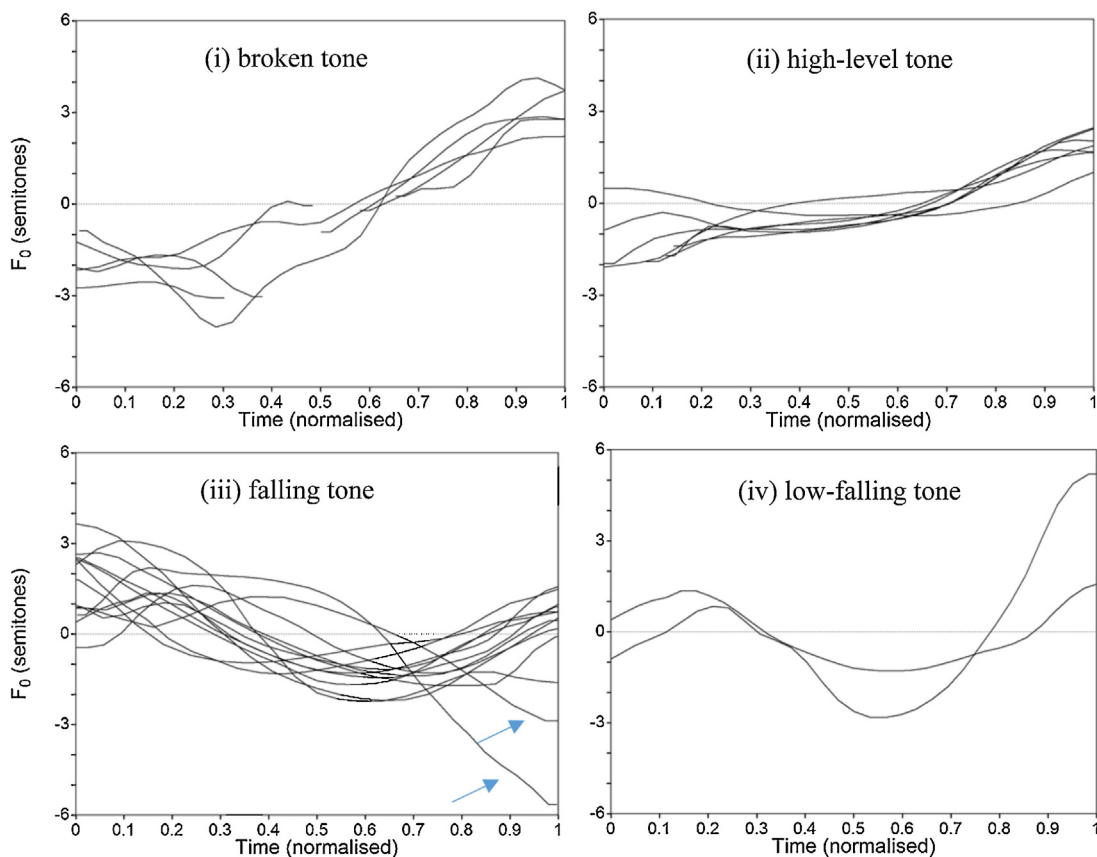

Fig. 5. Normalised fundamental frequency contours on final syllables in multi-word OIRs; the final syllables carry (i) the broken tone, $n = 5$, (ii) the high-level tone, $n = 7$, (iii) the falling tone, $n = 13$, and (iv) the low-falling tone, $n = 2$. The falling contours marked with arrows will be returned to later and analysed as deviant cases.

4.4. Deviant cases

Among the multi-word OIRs in the corpus, we found two utterances that do not have the rising contour found in the majority of tokens (see contours marked with arrows in Fig. 5). In extract (4) below, we present the example of the multi-word OIR *cái gì* produced with a falling pitch contour *(cái* is a classifier, lexically rising; *gì* is the question word 'what', lexically falling). It is an excerpt from a conversation between two friends, speakers M_QV and M_HH, who make an appointment to see each other on the following day. This extract illustrates the transition from the closing part in the telephone conversation, after the speakers have agreed on the date and time and are about to end the conversation (lines 1–3), to the part in which speaker M_QV asks his friend for his consent that the conversation be recorded and used for research purposes (from line 4). The request by M_QV starts in line 4 and is completed in line 6. Speaker M_HH gives his feedback signal *mm* meaning 'Yes' in lines 5 and 7, which does not indicate his understanding of what has been asked yet, but only his attention. This is clearer when he suddenly utters *Cái gì.* 'What!' with a falling pitch (line 9) signalling an understanding problem after speaker M_QV asks him explicitly whether he agrees to be recorded in line 7. This falling contour corresponds to the falling pitch of the lexical tone of the word *gì*. Speaker M_QV repeats the request again (lines 10–11) until M_HH understands the information (line 12) and then definitively gives his consent to be recorded in lines 14 and 16. The lack of a rise on the OIR in line 9, in combination with the sequential format of the OIR structure T−1 from line 6 to line 8, T0 in line 9, and T+1 from line 10 to line 11, indicates that speaker M_HH does not simply have a problem hearing what has been said. Instead, there is a referential or sequential mismatch between the closing of the telephone conversation and the request for his consent. This is evident from his reaction in line 12 'Oh yes, yes, got it!' in combination with the pauses between the lines in turn T+1 where M-QV repeats the request, and where M_HH would have been able to take the turn. This example shows that the lack of the rise on the OIR indicates that the problem that the speaker wants to convey that there has been a referential or sequential mismatch in conversation.

(4) Meeting ((M_QV, M_HH, aged between 57-59))

```
                  (…)
  1   M_HH    Ừ  được rồi
              Yes OK PART
              'Yes, ok then.'
  2   M_QV    Ừ   ừ
              'Yes yes.'
  3   M_HH    Nhớ.
              PART
              'OK then!'
  4   M_QV    Ừ thế   tôi ah nói thêm với ông cái      này
              Yes then I   tell further to you CLASS DEM.PRON
              'Yes, I want to tell you something.'
  5   M_HH    Mm.
              'Hm.'
              (0.7s)
  6   M_QV    Tức là    cái     cuộc thu   này        của tôi là tôi cũng dành để cho RELATIVE  nó        ┐
              Mean be CLASS recording DEM.PRON of me be I  also get   for    RELATIVE she        │
              nghiên cứu                                                                          │
              study                                                                               │  T-1
              'Namely, this recording I'm doing is for my relative's study.'                       │
              (0.5s)                                                                               │
  7   M_HH    Mm.                                                                                  │
              'Hm.'                                                                                │
  8   M_QV    Ông đồng ý chứ   hả?                                                                  ┘
              You agree  PART yes/no ques.
              'Do you agree?'
              (--)
->  9   M_HH  Cái gì.                                              ┐  T0
              'What!'                                              ┘
              (1.1s)
 10   M_QV    Tức là    cái    cuộc thu   này          này ((laughing))  ┐
              Mean be CLASS recording DEM.PRON DEM.PRON                  │
              'I mean the recording of our conversation.'                │
              (-)                                                        │
 11           Tức là   cuộc   nói chuyện  này          này               │  T+1
              Mean be CLASS conversation DEM.PRON DEM.PRON                │
              'I mean this conversation.'                                 │
              (0.5s)                                                      ┘
 12   M_HH    À    à    à ((loud)) rồi
```

```
                    Yes yes yes        PART
                    'Oh yes yes, got it!'
13 M_QV    Ờ   tức   là  nó  kết hợp   mà,
                    Yes mean be she combine PART
                    'Yes, I mean she combines (this task with her trip), you know.'
                    (.)
14 M_HH    Mm mm.
                    'Yes yes.'
15 M_QV    Ông đồng ý chứ    hả?
                    You agree  PART yes/no ques.
                    'Do you agree?'
16 M_HH    Nhất trí nhất trí
                    Agree    agree
                    'I agree I agree.'
17 M_QV    Ừ   nhất trí nhất trí ok ok.
                    Yes agree   agree  ok ok
                    'Yes, you agree you agree, OK, OK.'
```

The second token that is analysed as a deviant case is from the same conversation as extract (3) above, the conversation between speakers M_AK and F_CT regarding money for salaries. Here, speaker F_CT produces the utterance *Tiền gì.* 'Which money!' (*tiền* meaning 'money', *gì* being the question word 'what/which', both of which are lexically falling), again with a falling pitch contour. This part of the conversation is reproduced in extract (5):

(5) Colleague ((M_AK, F_CT are colleagues, aged between 30-45))

```
  1  M_AK    Cô  ơi   ah       tiền  tiền  lương tiền  gì  là có   rồi   đấy   tiền
                   Aunt PART hesitation money money salary money what be available PART PART money
                   gì   là  có   rồi   nhưng bây giờ cháu chưa xuống       được
                   what be available PART but    now    I    not go (down) able      ⌉
                   'Hi aunt, the money, money, the salary, some money is now available, but I cannot come to │ T-1
                   you (to give it to you).'                                                                  │
                   (1.4s)                                                                                     ⌋
-> 2  F_CT     Tiền   gì.       ⌉ T0
                   money what     ⌋
                   'Which money!'
  3  M_AK    Tiền    trả công cấy của các cô ấy  là có      rồi  nhớ đầu    tháng chín  ⌉
                   money pay work  of  PL. aunt 3.Ps. be available PART PART beginning September │
                   (Person name mang xuống nó gửi lại)                                          │ T+1
                   person name  bring down  he give back                                        │
                   'The money for paying for their work is now available, at the beginning of September │
                   (Colleague is going to bring you).'                                              ⌋
  4  F_CT     Ừ.
                   'Ok.'
  5  M_AK    Nhưng mà có  rút       nhiều không.  Rút      hết không?
                   But       CÓ withdraw much KHÔNG withdraw all KHÔNG
                   'But should I withdraw a large amount, should I withdraw all of it?'
                   (1.1s)
  6  F_CT     Ờ mày rút     hết cả cái    tiền  ấy   đi    chứ?
                   yes you withdraw all    CLASS money 3. Ps. PART PART
                   'Yes, withdraw all of it!'
```

Before this extract starts, the two speakers were talking about a number of pesticides that are still available at speaker F_CT's place. Then, in line 1, speaker M_AK starts a new topic, talking about the money to be paid to employees in the company. However, the information about the money concerned was not uttered clearly at the beginning of this turn, leading to a problem in understanding, accompanied by some degree of disbelief on the part of speaker F_CT in line 2. This is clear from the fairly long gap of 1.4 s between the trouble source turn, T−1 in line 1, and her OIR 'Which money!', T0, spoken with a falling intonation. In T+1 (line 3), speaker M_AK makes a repair elaborating on the type of money, i.e. money for employees, which is acknowledged by F_CT in line 4. However, the disbelief of F_CT, expressed by the lack of a rise on 'Which money!' in T0, is reinforced by the gaps (1.1 s long) in line 5 after the turn in which M_AK moves to the topic of withdrawing this money and before the turn in which F_CT agrees that this money can be withdrawn (line 6). Despite the fact that there may be other factors that influence the way the OIR in this extract is produced (e.g. F_CT might be annoyed about the way M_AT addresses the issue of money), we analysed this token as a deviant case.

To sum up, the two deviant cases analysed above show that OIRs which do not have the typical final rise in pitch indicate that besides the request for clarification or elaboration, the speakers also convey other meanings, such as

disbelief of what is being said by the other speaker, a referential or sequential mismatch in conversation. These cases support the finding that OIRs – requests for a repetition, clarification, elaboration or correction after mishearing or misunderstanding a previous utterance – are typically produced with a rising intonation, or at least with a final high pitch (depending on the identity of the lexical tone of the words involved). When this rise is not used, speakers convey additional meanings in their requests, such as referential or sequential mismatch or surprise/disbelief.

5. Interaction between repair intonation and lexical tones

5.1. Single-word other-initiation of repair

We analysed how the OIR intonation interacts with the lexical tones of the words *hả*, *gì*, *ai* and *ơi* in OIRs by comparing them to productions in the carrier sentence *Tôi dạy từ X cho bé* ''I teach the word X to the child'' (X = target word: *hả*, *gì*, *ai* or *ơi*). These words were identified as single-word OIRs in the telephone data of six speakers, speakers F_KT, F_TT, F_DL, F_DH. M_DA and M_QV, who were asked to read aloud only the words they used in telephone calls (three repetitions for each word). The target words were as follows:

hả (lexically low-falling, 'yes/no question word') read by speakers F_KT, F_TT, F_DL, M_DA, M_QV;
gì (lexically falling, 'what') read by speaker F_DH;
ai (lexically high-level, 'who') read by speakers F_DL and M_DA;
ơi (lexically high-level, 'vocative particle') read by speaker F_DH.

The target words were produced non-finally to preclude the influence of an intonational tone, for example, listing intonation, often used in word lists, so as to obtain pitch contours equivalent to the citation forms of the lexical tones, as reported in e.g. Kirby (2011).

The majority of the tokens elicited using read speech discussed in this section involve the word *hả* (lexically low-falling tone), since most OIRs used this word. We therefore illustrate this target word separately from the other three target words.

Fig. 6 shows that the pitch contours in *hả* OIRs (black lines) are higher than those in citation form (grey lines) for five speakers. They are higher either in terms of final pitch height (for speakers F_TT and F_DL) or pitch excursion (for speaker M_QV), or a combination of both (for speakers F_KT and M_DA). The identity of the lexical low-falling tone appears to be *obscured* by the rise used in OIRs in almost all cases, shown by the black lines, most of which entirely lack a fall. Although direct comparison across different recording situations is problematic, it is nonetheless interesting to note that in some cases, the OIR contours start at the same level as the lexical tone. This provides an indication that phonetic properties of the lexical tone may remain in the first portion of the signal for some speakers (F_TT and F_DL).

Fig. 7 illustrates the pitch contours produced on the utterance *gì*, lexically falling, and the utterances *ơi* and *ai*, lexically high-level. The pitch contours in OIRs are rising (for speakers F_DH and F_DL), or level (for speaker M_DA). As in the productions of *hả* by F_TT and F_DL, the OIR contours on *gì* and *ơi* produced by F_DH are rising during the second half of the syllable. Although the pitch contours produced on *ai* by M_DA show neither a high final pitch nor a large pitch excursion, it cannot be ruled out that other factors, such as the attitude of the speakers, the relationship between them (cf. Mạc et al., 2010), or possibly the sequential context in which the OIR occurs, may have influenced the pitch realisation in this case (see also section 4.1).

Taking the single-word OIRs of all other speakers into account, illustrated in Fig. 1 above, we can summarise that, regardless of the lexical tone of the words, a final rising intonation, or at least a non-falling intonation, is used to initiate a repair. For some speakers, the rising intonation is accompanied by a greater pitch excursion. For high-toned words, the OIR intonation may consist of a final rise but may also remain high level, as the lexical tone is already high. The rise or high pitch found in OIRs may be interpreted as a reflex of a final high boundary tone (H%), a tone associated with and loosely aligned with the end of the phrase. This captures the observation that the pitch contour is not always affected throughout the syllable, but is largely affected in the second half. The variability in the extent of the rise over the syllable can also be accounted for by such an analysis (Pierrehumbert and Beckman, 1988; Gussehoven, 2004; Ladd, 2008). This interaction between intonation and lexical tones has been reported in other tone languages, such as Thai (Luksaneeyanawin 1998) and Chengdu Chinese (Chang, 1958, see also Ladd, 2008, 156–168 for a summary).

5.2. Multi-word other-initiation of repair – supporting evidence

In this section we provide further supporting evidence for the high/rising pitch in OIRs. We compare syllables produced at the end of multi-word OIRs with syllables produced at the end of other utterances (e.g. an agreement) by the same
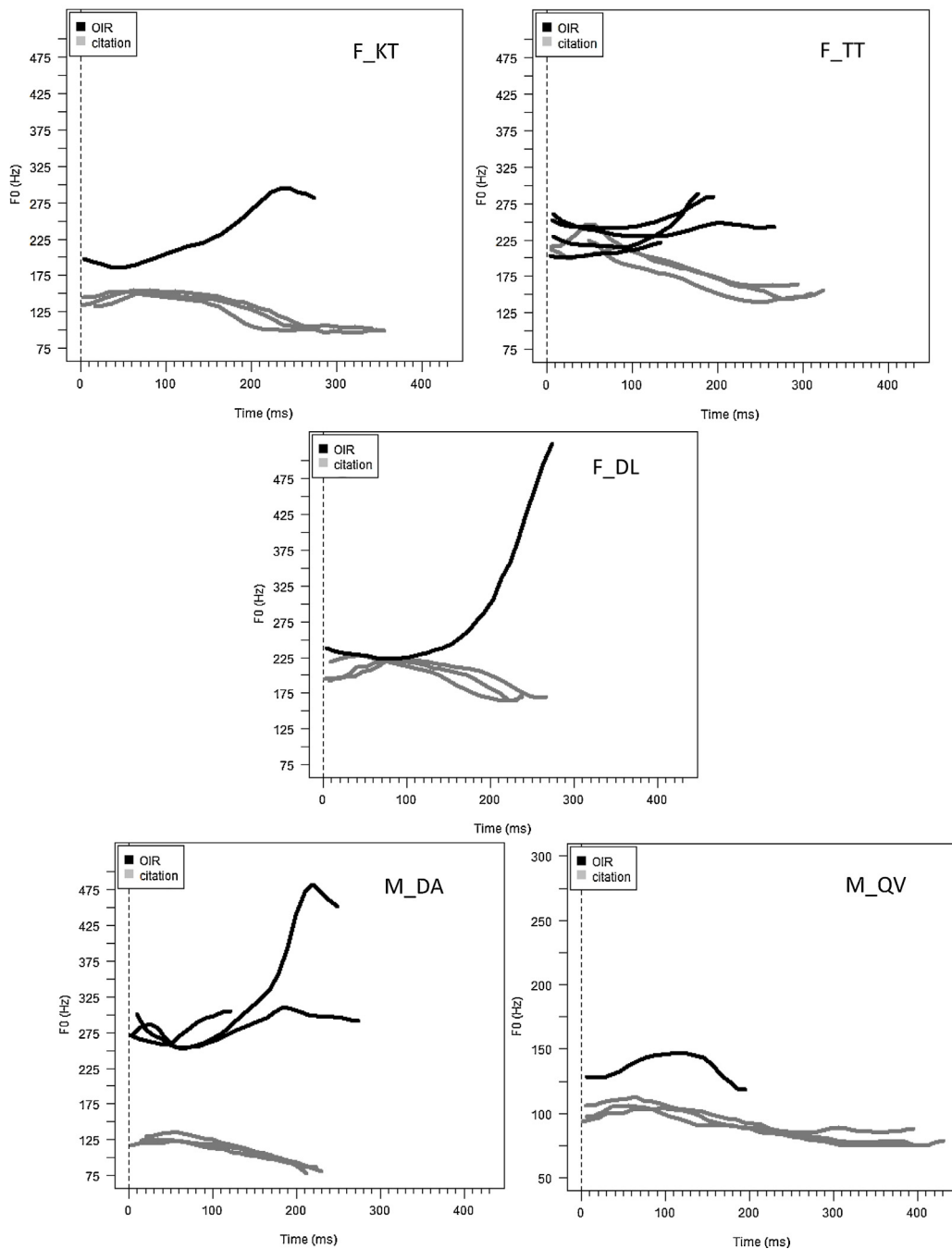
hả, lexically low-falling

Fig. 6. Fundamental frequency contours on the utterance/word *hả*, lexically low-falling, produced in OIRs (black lines) and citation form (grey lines, each speaker *n* = 3), by speakers F_KT, F_TT, F_DL, M_DA and M_QV.

speaker. The target syllables all involved the same lexical tone and have the same segmental makeup. Below we present two examples, one involving the lexical word *dòng* 'book line' and one involving the sentence particle *nữa* 'else'.

Fig. 8 illustrates two pitch contours produced on the syllable *dòng*, a lexical word with a falling lexical tone meaning 'book line', at the end of two utterances produced by speaker M_BDA. The first, shown in black, occurs in an OIR. The second, shown in grey, occurs when the speaker agrees to leave two empty lines in his homework.
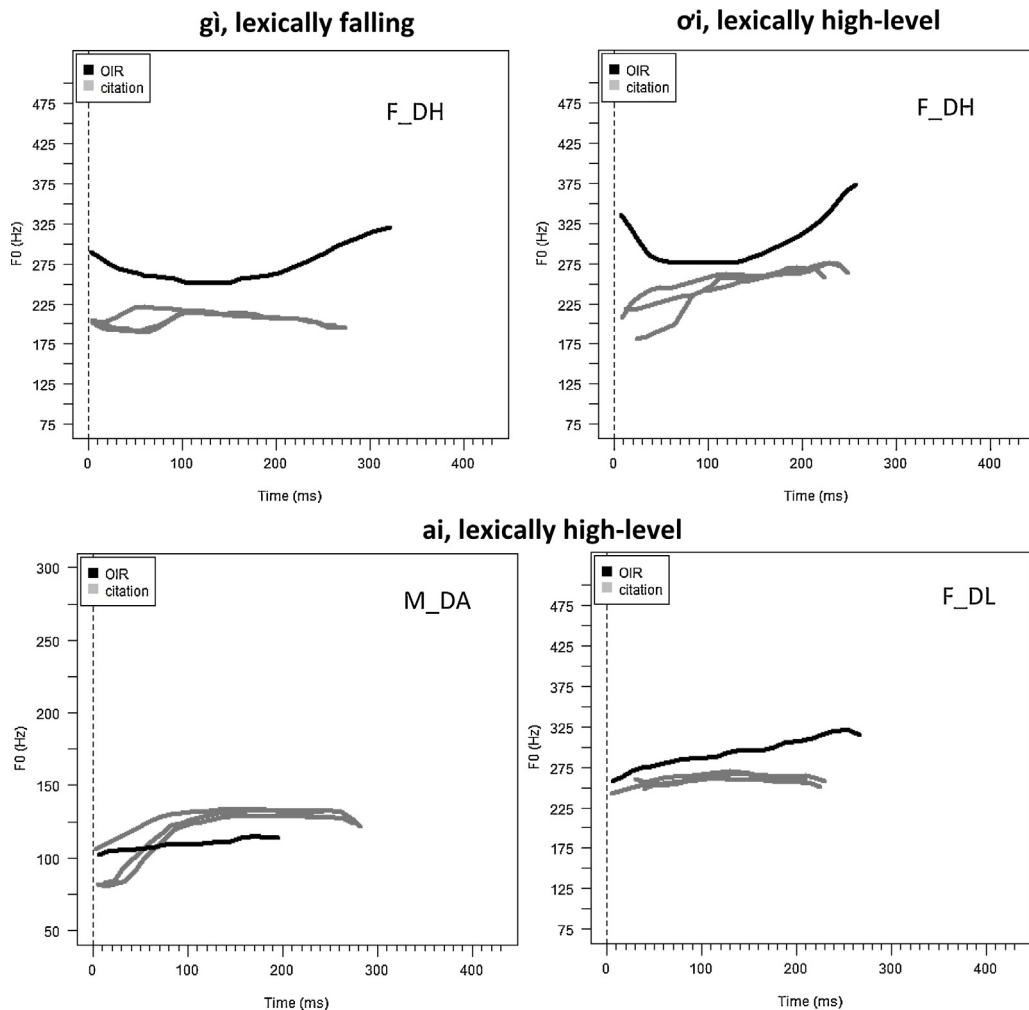
Fig. 7. Fundamental frequency contours on the utterances/words *gì*, lexically falling, *ơi* and *ai*, lexically high-level, produced in OIRs (black lines) and citation forms (grey lines, each speaker *n* = 3) by speakers F_DH, M_DA and F_DL.

In the OIR 'Leave how many lines?', the pitch contour on *dòng* has a final rise. For the utterance '[OK] Leave two lines.', where the speaker agrees on leaving two lines in the homework, the pitch falls, reflecting the lexical falling tone of the word. But in the OIR, the contour falls only until the last 25% of the signal, leaving room for a rising pitch used to mark an OIR. This is evident in the duration difference between the two utterances, the OIR being almost 40% longer than the agreement. This may be seen as an indication of syllable lengthening in order to accommodate the rise in pitch for the OIR.

Fig. 9 illustrates the pitch contours of the final three syllables of the utterances containing the particle *nữa* 'else', which has a lexical broken tone. These utterances were produced in a turn eliciting an agreement from an interlocutor (i) and in an OIR asking for clarification in (ii):

(i)   *Rồi thì lên cơ quan làm gì   nữa!*
       Then   go   office   do   what   else
       'Then it makes no sense to go to the office again!'
(ii)   *Lại học   gì   tiếp   nữa?*
       But study what further else
       'What else does (she) want to study?'

The differences between the realisations of *nữa* in (i) and (ii) in Fig. 9 involve the final pitch height, overall pitch excursion, voice quality and duration. There are some differences at the beginning of the final syllables in the two realisations due to progressive coarticulation from the lexical tone of the penultimate syllable (Brunelle, 2009a). *Gì* in (i)
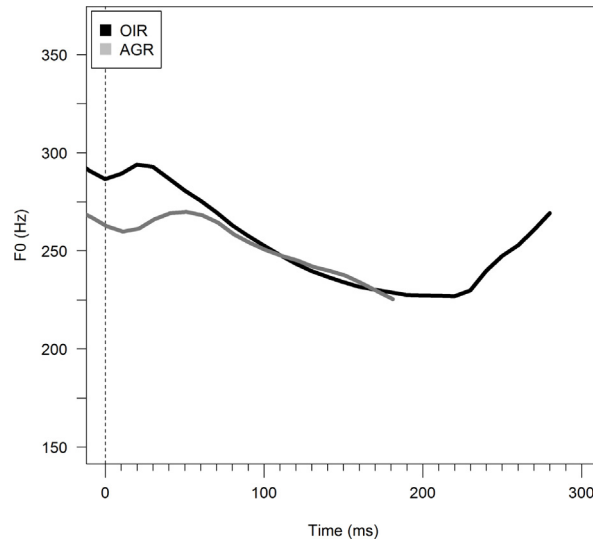
Fig. 8. Utterance-final syllable *dòng*, lexically falling, produced with a falling pitch contour in an agreement (AGR, grey) 'Leave two lines.', and the same word with a falling-rising pitch contour in an other-initiation of repair (OIR, black) 'Leave how many lines?' produced by speaker M_BDA.
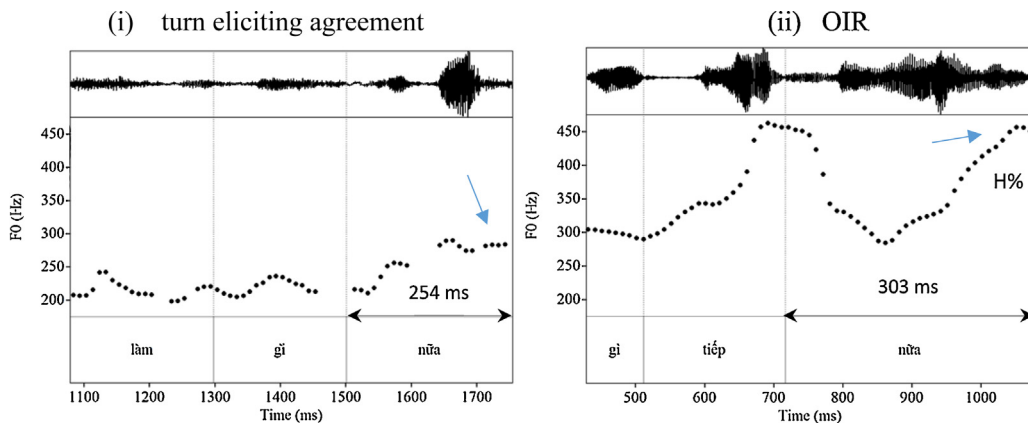

Fig. 9. The falling/low pitch (i) on final syllable *nữa*, lexically broken, in a turn eliciting an agreement from an interlocutor 'It makes no sense to go to the office again!' and the high rising pitch contour (ii) on *nữa* in an other-initiation of repair 'What else does she want to study?' produced by speaker F_TT.

has a falling tone and *tiếp* in (ii) has a rising tone, resulting in a higher initial F0 for *nữa* in (ii). Our focus is on the second half of the final syllables, as functions in discourse have been shown to affect the pitch contours here, both in this study and in previous work (Nguyễn et al., 2012). In fact, while in the utterance eliciting agreement from interlocutor (i) the contour ends with a low pitch, which is accompanied by creakiness/glottalisation, the contour in the OIR in (ii) rises to a very high pitch. The pitch excursion in the OIR is much higher for this speaker than it is in her turn eliciting agreement. Furthermore, while the glottalisation inherent in the lexical broken tone of the word is visible in the waveform and auditorily detectable in (i), it is minimal in (ii). This repair-related high pitch might interact with the glottalisation, a feature inherent in the lexical broken tone. It may be, for instance, that the rising part of the tone is exaggerated, reducing the need for voice quality cues. The duration of the OIR is clearly longer (303 ms vs. 254 ms).

The examples in Figs. 8 and 9 provide evidence that, even in multi-word utterances, a rise at the right edge of an utterance is used in other-initiations of repair, regardless of the lexical tone of the words. As in single word OIRs this final rise may be the phonetic exponent of a high boundary tone. In rising tone words (at least for the broken tone), the high boundary tone appears to make the lexical high/rising pitch even higher (this can sometimes be accompanied by increased loudness or a greater pitch excursion).

The interaction of the repair high boundary tone with the lexical tone, with which it overlaps at least in part, is further supported in the analysis presented here of multi-word utterances including both particles and lexical words. This analysis also shows that the final portion of the utterance – often the second half of the final syllable – is the domain in which intonation has an effect but does not extend to before the final syllable. Additionally, our examples show that syllables in OIRs tend to be lengthened, possibly to accommodate the high boundary tone used to mark OIRs (Ladd, 2008). Non-modal features of the lexical tones, in particular glottalisation, may be absent when there is an OIR intonation (e.g. the case of *nữa* in Fig. 9), but this is not always the case, and may be specific to particular lexical tones (see *dạ* in Fig. 2).[11]

6. Discussion and conclusions

The present study of Standard Vietnamese investigates the intonation used in single-word and multi-word other-initiations of repair, indicating a mishearing or a problem in understanding in conversational interaction. The qualitative analysis of the data shows that to carry out this action in discourse, speakers make predominant use of a *rising* pitch contour, an intonation pattern that has been found in a large number of languages (cf. Enfield et al., 2013; Dingemanse et al., 2013). These results correspond to reports in a number of previous studies on Vietnamese intonation that interrogatives are marked by a final rise (Đỗ et al., 1998; Nguyễn and Boulakia, 1999; Nguyễn, 2004; Vũ et al., 2006; Brunelle et al., 2012). This rising intonation has been found in the majority of OIRs in the corpus, regardless of part of speech or lexical tone. In OIRs with words containing the lexical high-level tone, a level contour was often used, possibly because the lexical tone is already high. The rising intonation in OIRs can be analysed as the phonetic reflex of a high boundary tone (H%) affecting the realisation of the lexical tone of the final syllable of the utterance. In a very small number of cases falling pitch was found. We analysed these as deviant cases, since they appeared to indicate a referential/sequential mismatch or disbelief rather than a simple problem of mishearing.

In single-word utterances, which included only function words and particles, the interaction between the intonational rise and the lexical tone can be analysed in terms of variable overlap of lexical tone with the final intonational high tone, along the lines of gestural overlap in Articulatory Phonology (Goldstein and Fowler, 2003; Katsika et al., 2014). The intonational tone can *overlap partially* with the lexical tone, in particular in the second half of the syllable, or the intonational tone can *overlap completely*. In the case of low-toned words, this overlap can leave little or no trace of the lexical tone in the signal. The same rise in pitch could be found on the final syllable of multi-word utterances, indicating not only that the pitch of lexical words can be affected by the intonational tone, but also that the effect was confined to the final syllable.

We have proposed that the local final pitch found in OIRs can be modelled using intonational tones, i.e. a high boundary tone is used to initiate a repair in the interactional context investigated here (cf. Ladd, 2008, section 4.3). Although Brunelle et al. (2012) showed that boundary tones were not consistently used by all speakers in laboratory speech when expressing pragmatic functions in Standard Vietnamese, such as in distinguishing statements from questions, results here indicate that they are used more consistently in OIRs in spontaneous conversation. Together with our studies on the low, falling intonation of backchannels, we thus provide evidence indicating that in this language intonational tones are used consistently in talk-in-interaction (Hạ, 2010; Hạ and Grice, 2010; Hạ, 2012). Moreover, the low, falling intonation, analysed as a L% boundary tone, is apparent in backchannels even in written discourse, as observations of digital discourse such as Facebook posts indicate. For instance, the word *vâng*, which has a lexical high-level tone (no tone diacritics), is sometimes written in the backchannel context indicating agreement or confirmation as *vầng*, with the diacritic of the lexical low tone. The fact that native speakers represent intonational tones as lexical tones in orthographic writing supports the view that intonation in Vietnamese can be interpreted in terms of discrete tones. It is thus possible, given the above, that intonational tones play – or at least have begun to play – a greater role in Vietnamese discourse than descriptions of the language have so far assumed.

This paper suggests that prosody in interaction is a fertile research area for studies on the interaction between lexical tone and intonation in tone languages, and shows that, for Vietnamese, repair is an action in conversation in which the prosodically conveyed pragmatic expression in interactional discourse could be seen as more relevant than the lexical tones of the words. Follow-up studies will need to extend the number of target tokens analysed, so as to statistically test the generalizability of these results, and in particular the relevance of the rising intonation for the interpretation of an utterance as an OIR. Despite the fact that Vietnamese is a language that has been argued to make little use of intonation due to the extensive use of sentence particles, it appears that there are at least some interactional contexts in which intonational tones play a role in conveying pragmatic meaning.

---

[11] The low glottalised tone *nặng*, e.g. in *dạ*, has been attested as having a ''dramatic glottal stop'', while the broken tone *ngã*, e.g. in *nữa*, only shows ''glottal constriction'' (Brunelle, 2015, cf. Kirby, 2011). The dramatic glottal stop in *dạ* might be the reason why the non-modal feature of glottalisation is more resistant to change in *dạ* than it is in *nữa* when these words are used to initiate a repair.

Acknowledgements

Appendix A. Abbreviations and transcription conventions

Abbreviations

NEG	negation
CLASS	classifier
PL	plural
PART	particle
DEM.PRON	demonstrative pronoun

Transcription conventions

RELATIVE	replacement of person information for the purpose of anonymity
(...)	previous part of conversation
( )	unclear talk
(( ))	notions that help establish the dialogue context
(-) (--) (---)	short pause approx. 0.2–0.5 s, long pause approx. 0.5–0.8 s, longer pause approx. 0.8–1.0 s
(1.1 s)	pauses above 1 s
(.)	micro pause
[	speech overlap
[	
<<f> >	forte, loud

Appendix B. List of particles occurring in investigated excerpts

ý or *(đ)ấy*	particle used for third person, e.g. *anh ấy* 'he', *chị ấy* 'she', in extract (1) *trường (â)ý* 'that school', recognitional particle
ơi	vocative
rồi	'already' or expression of a situation having occurred
đấy	final particle, often occurring in questions in casual speech
nhớ/nhá	final particles, often occurring in statements or imperatives in casual speech
có...không	yes/no question frame
không	yes/no question final particle
đi (chứ)	imperative particle(s)
thì	focus particle
mà	final particle expressing an explanation, an emphasis or confirming a fact
chứ gì	question particles 'right?' or 'correct?'
chứ	particle, often occurs in combination with other question particles, e.g. *chứ hả* 'polar question'

Appendix C. List of multi-word other-initiations of repair investigated

In total, 27 multi-word OIRs were analysed, including 2 deviant cases:
<u>Check for understanding</u>

1. *Thủy Định hả?*
   Thuy Dinh ques.word
   Thuy Dinh? (names of a couple)

Clarification of a co-referent

2. *Vở   học thêm   ý   gì?*
   Book   tutoring class   part.   what
   The book for the private class?

3. *Vào chỗ   chồng   cái Đức   ý   gì?*
   Go   place   husband   classifier Duc   part.   what
   (He plans to) apply for the position at the place where Duc's husband works?

4. *Con   Huệ là năm tám   mốt ý   gì?*
   Classifier   Hue   be   year   eighty   one   part.   what
   Hue was born in 1981?

5. *Lại   học gì   tiếp   nữa?*
   Again   learn   what   further   else
   What else does she want to study?

Request for elaboration

6. *Gọi   làm gì?*
   Call   do   what
   Why did he call?

7. *Thì ở dưới   đấy   là gì?*
   So   in   underneath   there   be   what
   (ok) what comes after that?

8. *Ở dưới   cái   đấy   là gì?*
   In   underneath   classifier   that/there   be   what
   So what comes after that?

9. *Cách   bao nhiêu dòng?*
   Leave   how   many   line
   How many lines do I have to leave?

10. *Số   di động hay số   nhà?*
    Number   mobile   or   number   home
    Mobile phone or landline number?

11. *Làm   sao?*
    Do   how
    What comes next?

12. *Gì nữa? (three times)*
    What else?

15. *Tổng kết   sách giáo khoa gì   nữa?*
    Summarise   book   text   what   else
    Summarise the contents in the textbook, then what else?

Mishearing

16. *Ý   gì   cơ?*
    Point   what   part.
    Which point?

17. *Tìm hiểu chi tiết gì?*
    Study detail what
    Study what detail?

18. *Gì cơ? (three times)*
    What part.
    What?

21. *Ý nghĩa của các gì?*
    Meaning of PL. what
    The meaning of what?

22. *Tô đậm gì cơ?*
    Emphasise what part.
    Emphasise what?

23. *Ghi hoàn cảnh của ai?*
    Write situation of who
    Write down the situation of whom?

24. *Hình như cái gì?*
    Probably classifier what
    Probably what?

Problem in understanding

25. *Sao lại phải cách hai dòng?*
    Why should leave two line
    Why do I need to leave two lines?

Two deviant cases (disbelief or referential problem)

26. *Tiền gì!*
    Money what
    Which money!

27. *Cái gì!*
    Classifier what
    What!

References

Alves, Mark, 1997. Tonal features and the development of Vietnamese tones. Working Papers in Linguistics, vol. 27. Department of University of Hawaii at Manoapp. 1–13.

Boersma, Paul, Weenink, David, 2015. Praat: Doing Phonetics by Computer. Retrieved from http://www.praat.org/

Brunelle, Marc, 2009a. Northern and Southern Vietnamese tone coarticulation: a comparative case study. J. Southeast Asian Linguist. Soc. 1, 49–62.

Brunelle, Marc, 2009b. Tone perception in Northern and Southern Vietnamese. J. Phon. 37, 79–96.

Brunelle, Marc, 2015. Vietnamese. In: Jenny, Mathias, Sidwell, Paul (Eds.), A Handbook of Austroasiatic Languages, vol. 2. Brill, Leiden/Boston, pp. 909–953.

Brunelle, Marc, Hạ, Kiều Phương, Grice, Martine, 2012. Intonation in Northern Vietnamese. Linguist. Rev. 29 (1), 3–36.

Cangemi, Francesco, 2015. Mausmooth. Retrievable online at http://phonetik.phil-fak.uni-koeln.de/fcangemi.html

Chang, Nien-Chuang T., 1958. Tones and intonation in Chengtu dialect (Szechuan, China). Phonetica 2, 59–85.

Cheung, Kwan-Hin, 1986. The phonology of present-day Cantonese Doctoral dissertation. University College London, London.

Couper-Kuhlen, Elisabeth, 1992. Contextualizing discourse: the prosody of interactive repair. In: Auer, Peter, di Luzio, Aldo (Eds.), The Contextualization of Language. John Benjamins, Amsterdam/Philadelphia, pp. 337–364.

Dehé, Nicole, 2015. The intonation of the Icelandic other-initiated repair expressions Ha 'Huh' and Hvað segirðu/Hvað sagðirðu 'What do/did you say'. Nordic J. Linguist. 38 (2), 189–219.

Dingemanse, Mark, Torreira, Francisco, Enfield, Nick J., 2013. Is "Huh?" A universal word? Conversational infrastructure and the convergent evolution of linguistic items. PLOS ONE 8 (11), e78273. http://dx.doi.org/10.1371/journal.pone.0078273

Dingemanse, Mark, Roberts, Seán G., Baranova, Julija, Blythe, Joe, Drew, Paul, Floyd, Simeon, Gisladottir, Rosa S., Kendrick, Kobin H., Levinson, Stephen C., Manrique, Elisabeth, Rossi, Giovanni, Enfield, Nick J., 2015. Universal principles in the repair of communication problems. PLOS ONE 10 (9), e0136100. http://dx.doi.org/10.1371/journal.pone.0136100

Đỗ, Thế Dũng, Trần, Thiên Hương, Boulakia, Georges, 1998. Intonation in Vietnamese. In: Hirst, Daniel, Di Cristo, Albert (Eds.), Intonation Systems: A Survey of Twenty Languages. Cambridge University Press, Cambridge, pp. 395–416.

Egbert, Maria, Golato, Andrea, Robinson, Jeffrey D., 2009. Repairing reference. In: Sidnell, Jack (Ed.), Conversation Analysis: Comparative Perspectives. Cambridge University Press, pp. 104–132.

Enfield, Nick J., Dingemanse, Mark, Baranova, Julija, Blythe, Joe, Brown, Penelope, Dirksmeyer, Tyko, Drew, Paul, Floyd, Simeon, Gipper, Sonja, Gisladottir, Rosa S., Hoymann, Gertie, Kendrick, Kobin H., Levinson, Stephen C., Magyari, Lilla, Manrique, Elisabeth, Rossi, Giovanni, San Roque, Lila, Torreira, Francisco, 2013. Huh? What? – A first survey in 21 languages. In: Hayashi, Makoto, Raymond, Geoffrey, Sidnell, Jack (Eds.), Conversational Repair, Human Understanding. Cambridge University Press, Cambridge, pp. 343–380.

Goldstein, Louis, Fowler, Carol A., 2003. Articulatory phonology: a phonology for public language use. In: Schiller, Niels O., Meyer, Antje S. (Eds.), Phonetics and Phonology in Language Comprehension and Production. Mouton de Gruyter, pp. 159–207.

Gumperz, John J., 1992. Contextualization revisited. In: Auer, Peter, di Luzio, Aldo (Eds.), The Contextualization of Language. John Benjamins, Amsterdam/Philadelphia, pp. 39–53.

Gussehoven, Carlos, 2004. The Phonology of Tone and Intonation. CUP, Cambridge.

Hạ, Kiều Phương, 2010. Prosody of Vietnamese from an interactional perspective: ờ, ừ and vâng in backchannels and requests for information. J. Southeast Asian Linguist. Soc. 3 (1), 56–76.

Hạ, Kiều Phương, 2012. Prosody in Vietnamese – Intonational form and function of short utterances in conversation PhD thesis. In: Asia-Pacific Linguistics 002 (SEAMLES 001, open access). The Australian National University, Canberra.

Hạ, Kiều Phương, 2013. Prosodic means in repair initiation as an activity in Northern Vietnamese conversation. In: Hole, Daniel, Löbel, Elisabeth (Eds.), Linguistics of Vietnamese – An International Survey. De Gruyter, Berlin, New York, pp. 35–54.

Hạ, Kiều Phương, Grice, Martine, 2010. Modelling the interaction of intonation and lexical tone in Vietnamese. In: Proceedings of the 5th International Conference on Speech Prosody 2010, pp. 1–4.

Hayashi, Makoto, Raymond, Geoffrey, Sidnell, Jack (Eds.), 2013. Conversational repair and human understanding. Studies in Interactional Sociolinguistics, vol. 30. Cambridge University Press, New York.

Honda, Koichi, 2006. F0 and phonation types in Nghe Tinh Vietnamese tones. In: Warren, Paul, Watson, Catherine I. (Eds.), Proceedings of the 11th Australian International Conference on Speech Sciences, pp. 454–459.

Katsika, Argyro, Krivokapić, Jelena, Mooshammer, Christine, Tiede, Mark, Goldstein, Louis, 2014. The coordination of boundary tones and their interaction with prominence. J. Phon. 44, 62–82.

Kirby, James, 2011. Vietnamese (Hanoi Vietnamese). J. Int. Phon. Assoc. 41 (3), 381–392.

Ladd, Robert D., 2008. Intonational Phonology, 2nd ed. Cambridge University Press, Cambridge.

Mạc, Đăng Khoa, Aubergé, Véronique, Rilliard, Albert, Castelli, Eric, 2010. Cross-cultural perception of Vietnamese audio–visual prosodic attitudes. In: Proceedings of the 5th International Conference on Speech Prosody 2010.

Michaud, Alexis, 2004. Final consonants and glottalization: new perspectives from Hanoi Vietnamese. Phonetica 61, 119–146.

Nguyễn, Thị Thanh Hoa, 2004. Contribution à l'étude de la prosodie du Vietnamien: Variations de l'intonation dans les modalités – assertive, interrogative et imperative Thèse de doctorat en Lettres, sciences humaines et sociales, Paris 7.

Nguyễn, Thị Thanh Hoa, Boulakia, Georges, 1999. Another look at Vietnamese intonation. In: Proceedings of the 14th International Congress of Phonetic Sciences, pp. 2399–2402.

Nguyễn, Văn Lợi, Edmondson, Jerold A., 1998. Tones and voice quality in modern northern Vietnamese: instrumental case studies. Mon-Khmer Stud. 28, 1–18.

Nguyễn, Thị Lan, Michaud, Alexis, Trần, Đỗ Đạt, 2012. Tonal coarticulation on particles in Vietnamese language. In: International Conference on Asian Language Processing, pp. 221–224.

Nguyễn, Thị Lan, Michaud, Alexis, Trần, Đỗ Đạt, Mạc, Đăng Khoa, 2013. The interplay of intonation and complex lexical tones: how speaker attitudes affect the realization of glottalization on Vietnamese sentence-final particles. Interspeech 2013.

Nolan, Francis, 2003. Intonational equivalence: an experimental evaluation of pitch scales. In: Proceedings of the 15th International Congress of Phonetic Sciences, pp. 771–774.

Phạm, Andrea H., 2003. Vietnamese Tone. A New Analysis. Routledge, New York.

Pierrehumbert, Janet, Beckman, Mary E., 1988. Japanese Tone Structure. In: Linguistic Inquiry Monograph 15. MIT Press, Cambridge.

Pittayaporn, Pittayawat, August 5, 2007 Prosody of final particles in Thai: interaction between lexical tones and intonation. In: Paper presented at the Workshop on Intonational Phonology: Understudied or Fieldwork Languages. Saarland University, Saarbrücken, Germany.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria Available online at: https://www.R-project.org

Sacks, Harvey, Schegloff, Emanuel A., Jefferson, Gail, 1974. A simplest systematics for the organization of turn-taking for conversation. Language 50 (4), 696–735.

Schegloff, Emanuel A., 1997. Practices and actions: boundary cases of other-initiated repair. Discourse Process. 23 (3), 499–545.

Schegloff, Emanuel A., Jefferson, Gail, Sacks, Harvey, 1977. The preference for self-correction in the organization of repair in conversation. Language 53, 361–382.

Selkirk, Elisabeth, 1996. The prosodic structure of function words. In: Beckman, Jill N., Dickey, Laura W., Urbanczyk, Suzanne (Eds.), Papers in Optimality Theory. GLSA Publications, Amherst, MA, pp. 439–470.

Selting, Margret, 1996. Prosody as an activity-type distinctive cue in conversation: the case of the so-called 'astonished' questions in repair initiation. In: Couper-Kuhlen, Elisabeth, Selting, Margret (Eds.), Prosody in Conversation. Interactional Studies. Studies in Interactional Sociolinguistics, vol. 12. Cambridge University Press, pp. 231–270.

Ten Have, Paul, 2007. Doing Conversation Analysis. A Practical Guide, 2nd ed. Sage Publications.

Thompson, Laurence, 1965. Vietnamese Reference Grammar. University of Washington Press, Seattle.

Trần, Hương Mai, 1967. Tones and intonation in South Vietnamese. In: Nguyễn, Đăng Liêm, Mai, Trần Hương, Dellinger, David (Eds.), Series A – Occasional Papers #9, Papers in Southeast Asian Linguistics No. 1. Linguistics Circle of Canberra, Canberra.

Vũ, Minh Quang, Trần, Đỗ Đạt, Castelli, Eric, 2006. Prosody of interrogative and affirmative sentences in Vietnamese language: analysis and perceptive results. In: Interspeech 2006, .

Wu, Ruey-Jiuan Regina, 2006. Initiating repair and beyond: the use of two repeat-formatted repair initiations in Mandarin conversation. Discourse Process. 41 (1), 69–109.

Wu, Wing-Li, 2008. An acoustic phonetic study of the intonation of sentence-final particles in Hong Kong Cantonese. Asian Soc. Sci. 4 (2), 23–29.