Cantonese turn-initial minimal particles: annotation of discourse-interactional functions in dialog corpora

Andreas Liesenfeld
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
lies0002@ntu.edu.sg

Abstract

This interactional linguistic study is concerned with the annotation of discourse-interactional functions of turn-initial particles in Cantonese conversation. These particles (or intersections) are commonly transcribed as *ngo* (哦), *ng* (嗯), *aa* (啊), *aak* (呃) and can format a range of functions both as turn-initial utterances or as stand-alone turns. Based on the analysis of 20 hours of naturally-occurring video corpus data, the study identifies five discourse-interactional functions that the most 'minimal' (i.e. shortest and mostly monophthongic) of these utterances can format: continuers, positive response tokens, change-of-state tokens, turn management tokens and repair initiators. I then show that three dimensions have to be taken into account to annotate those functions: sequential position, pitch contour and phonetic production format. In contrast to existing annotation taxonomies that directly map production format to function, I argue that discourse-interactional functions of these particles can only be annotated with reasonable accuracy if at least these three structural dimensions are taken into account. I conclude with discussing the relation between sequential position, sound and pitch format for each function.

1 Introduction

Turn-initial particles are short utterances such as *oh*, *huh* or *mmhm* in English that appear in turn-initial position and that can stand alone as turns. These particles have important functions in the joint construction of conversation and can constitute various actions depending on their sequential environment and production format. This study examines turn-initial particles in naturally-occurring Cantonese conversation, explicates some of their discourse-interactional functions, and examines the relation of their interactional uses to some aspects of their phonetic and prosodic production formats.

The utterances under study are a range of particles that occur in turn-initial or turn-constructional unit (TCU)-initial position and that are commonly described as particles or interjections (嘆詞) and transcribed using Chinese characters such as *ngo* (哦), *aa* (啊), *ng* (嗯), *aak* (呃). Specifically, the study focuses on the most 'minimal' turn-initial particles, those that are formatted using monophthongic and nasal utterances. It is not concerned with other 'larger' particles such as *ei* (誒), *ai yo* (唉唷) or *ji aa* (咦呀).

Figure 1 illustrates that such particles are not only a common phenomenon in naturally-occurring conversation, but also that these minimal utterances can format a range of different discourse-interactional functions depending on their sequential position and specific production formats. One of the speakers (P2) here produces three such particles in only a couple of seconds of talk that each format a different discourse-interactional function, a change-of-state token (Line 04), a continuer (Line 06) and a repair initiator (Line 08).

The accurate annotation of these particles is an integral part of any larger dialog act taxonomy that aims to further process speech act formation or model speaker intent. Notably, in Figure 1, each of the minimal utterances features a different produc-

Pacific Asia Conference on Language, Information and Computation (PACLIC 33), pages 471-479, Hakodate, Japan, September 13-15, 2019

Figure 1: Corpus excerpt showing three different discourse-interactional functions

Data excerpt (MYCANCOR 022 (04:53-05:06)) from a conversation between two participants, P1 and P2. Previous to the beginning of this excerpt, the topic of P1's partner was brought up and P2 inquired how P1 met her partner.

01 P2 form four ge3 si4 hau6 sik1(.) keoi5 tung4 nei5 tung4 hok6
form four嘅時候識佢同你同學
form four PAR time know 3SG with 2SG classmate
Got to know ((your partner)) in form four? Was he studying with you?

02 P1 m4 jat1 joeng6(.)ngo5 ngo5 dei2 hai6 jan1 wai4 tung4 jat1 go3 lou5 si1 maa3
唔一樣我我哋係因爲同一個老師嘛
NEG same 1SG 1PL be because same one CL teacher PAR
It wasn't like this, I, we were, because of the same teacher.

03 P1 go3 daa2 ngok6 tyun4 [>go3 lou5 si1<
個打樂團 個老師
CL band CL teacher
That band... teacher.

04⇒P2 [↑ngo::
哦
INT
Oh.

05 P1 jin4 zi1 hau6 ngo5 dei2 hok6 haau6 heoi3 keoi5 dei2 daai6 hok6 gaau1 lau4
然之後 我哋學校去佢哋大學交流
then 1PL school go 3PL university exchange
And then our school had an exchange with their university.

06⇒P2 ↓aa1.
啊
INT
Okay.

07 P1 zi1 hau6 hai6 go2 dou6 zau6 sik1 zo2 keoi5 lo3
之後係嗰度就識咗佢咯
then be there just know MOR 3SG PAR
Then I got to know him over there.

08⇒P2 ↑aa1 go2 jat1 baai3 ze1
啊[嗰一拜嗌
INT that one week PAR
Huh, in just a week?

09 P1 [hai6
係
is
Yes.

10 P1 |jat1 baai3
一拜
one week
A week.
|((P1 nods))

tion format. The change-of-state token in line 04 that displays a change to a state of knowing or understanding is formatted with [↑ngo::] (featuring a rising pitch contour). The continuer in line 06 is formatted as [↓aa1] (featuring a falling pitch contour), and the repair initiator in line 08 is also formatted with [↑aa1], this time featuring a rising pitch contour. Given this discrepancy in production format, can the discourse functions of these utterances be annotated by analysing their prosodic-phonetic form alone? I argue that, while production format is an important factor in the constitution of different functions, additional structural dimensions have to be taken into account to annotate these utterances accurately. Existing approaches to the functional annotation of these minimal yet important monophthongic and nasal utterances largely focus on analysing their production format, especially pitch contour, and propose to annotate functions directly mapped to specific production formats. Before discussing the results of the annotation efforts, I briefly review related work on particles and existing annotation taxonomies for Chinese.

1.1 Related work

Turn-initial particles are well-studied in both Mandarin and Cantonese and various reference grammars of spoken Chinese have described these utterances, referring to them as particles, interjections or non-lexical utterances (for Mandarin see, for instance, Chao (1965), Hu (1987) Li and Thompson (1989) and for Cantonese Killingley (1993), Cheung (2007), Matthews and Yip (2013)).

Only few studies, however, focus on turn-initial particles in naturally-occurring talk-in-interaction in particular. Studies coming out of interactional linguistics and conversation analysis have mainly focused on Mandarin Chinese, but nonetheless provide important insights in the work that turn-initial particles do in both languages. These studies have described a range of functions that turn-initial particles are involved in, mostly exploring a specific action that may be formatted using minimal particles.

Xudong (2008) examines continuers in Mandarin conversation and describes several uses of turn-initial particles under the topic of "listener responses". Oralova (2016) examines "minimal response tokens" in Mandarin and, focusing on *en* (嗯), shows that this particle can format continuers and positive response tokens. Also focusing on *en* (嗯) in turn-initial position, Xu (2009) describes "resumptive openers" in Mandarin.

Existing literature on repair in Mandarin also deals with turn-initial particles. Wu (2006) and Tseng (2013) show how various minimal particles including *en* (嗯) and *aa* (啊) can format repair initiators in Mandarin.

To the author's knowledge, no previous work on turn-initial particles in Cantonese has been done in the fields of interactional linguistics or conversation analysis.

Besides the above interactional linguistic studies that mainly explore the formation of a specific action in talk-in-interaction, a comprehensive annotation guideline for particles in Mandarin Chinese speech has been developed by Ping et al. (2014) (see also Lee et al. 2017). The study provides a detailed guide of how to annotate discourse-interactional functions of turn-initial particles based on the examination of production format and pitch contour. By mapping a specific format to one or more functions, particles can this way be annotated with reasonable accuracy. For instance, all three turn-initial particles previously shown in Figure 1 can be distinguished by examining their production format. However, this 'form-to-function mapping' approach is not well-suited to annotate the use of similar production formats that constitutes different functions in different sequential environments. Consider Figure 2 that illustrates this limitation.

Here two participants use a turn-initial particle of similar production format "*aa1*" (啊), but the utterance formats a positive response token in Data Excerpt 1 (Line 03) and a repair initiator in Data Excerpt 2 (Line 02). Notably, the participants do not treat this 'ambiguity' as problematic and follow up with a turn that displays action ascription respectively. This illustrates that, in order to annotate different functions of similar particles, another structural dimension has to be taken into account: sequential position.

2 Methodology

This study is situated in the field of interactional linguistics (Couper-Kuhlen and Selting, 2018) and

Figure 2: Example of use of 啊 (*aa1*) as positive response token and as repair initiator

```
Data excerpt (1): Positive response token 啊 (aa1)
MYCANCOR 021  (04:24-04:29)

01  P1   kei4(.)keoi5 hai6 mi1 ceot1 hoi1 ceot1 hoi1 kei4 taa1 dei6 fong1.
         其 佢係咪出開 出開其它地方
         actually 2SG is NEG go out go out other place
         Actu ((ally)), didn't she leave, leave ((that workplace)) for another place.

02  P2   >bin1 go3< ji3 mai5?
         邊個 薏米
         who person name
         Who, Barley.

03⇒P1    aa1:
          啊
          INT
          Uh.

04  P2   m4 zi1 ak1.
         唔知喔
         NEG know PAR
         I don't know.



Data excerpt (2): Repair initiator 啊 (aa1)
MYCANCOR 009  (00:45-00:52)

01  P1   nei5 dei2 zung6 jau5 cyun4 bin1 go3.
         你哋仲有傳邊個
         2SG also have convey which one
         Who else did you guys talk about?

02⇒P2    aa1:
          啊 (0.2)
          INT
          Huh?

03  P1   ceoi4 zo2 aa3 ciu1 zung6 jau5 cyun4 bin1 go3.
         除咗阿超仲有傳邊個
         despite Ah-Ciu also have convey which one
         Despite Ah-Ciu who else have you guys been talking about?

04  P2   mou5 aa1.
         冇啊
         not have PAR
         Nobody.
```

examines the use of minimal (monophthongic and nasal) utterances in turn-initial position as part of processes of action formation and ascription in talk-in-interaction (Levinson, 2013).

I present the results of a manual annotation of 20 hours of corpus data from a video corpus of naturally-occurring everyday talk. Around 484 instances of the production of turn-initial particles using monophthongic and nasal utterances were annotated in the corpus. The data was annotated according to interactional linguistic principles, employing the next-turn proof procedure to distinguish five discourse-interactional functions that participants commonly format using monophthongic (and nasal) utterances: continuers, positive response tokens, change-of-state tokens, turn-management tokens and repair initiators. All utterances were then transcribed using the International Phonetic Alphabet (IPA) for Cantonese (Zee, 1991) and a pitch contour analysis was conducted.

All data excerpts are from the 'MYCanCor' corpus of colloquial Malaysian Cantonese, a video corpus of 20 hours of naturally-occurring everyday conversation (Liesenfeld, 2018). The corpus data is transcribed in accordance with common practice in the field of interactional linguistics, using a four-line format consisting of Jyutping romanisation, Chinese characters including the Hong Kong Supplementary Character Set (HKSCS), word-by-word translation and English translation. This is a corpus of Cantonese Chinese as spoken in contemporary Malaysia. While their are differences between this variety of Cantonese and, for instance, Cantonese spoken in Hong Kong, the authors expect that the findings presented in this paper with regards to turn-initial particles are applicable across different Cantonese speech communities.

3 Results

The question that this study addresses is what discourse-interactional functions do participants format when uttering turn-initial minimal particles and what production formats do they commonly use to produce them. The aim is to examine the relationship between sequential position and the prosodic-phonetic properties of these utterances, and, by doing so, to contribute to a better understanding of how to annotate discourse-interactional functions of these particles in colloquial conversation.

For each of the five functions identified in the data set, I show an overview of their smoothed pitch contour and phonetic transcription based on IPA, and briefly discuss the relationship between the two properties.

3.1 Continuers

Continuers (or receipt tokens) are utterances that format a continuer action, i.e. that invite an interlocutor to go on talking. These utterances commonly appear at transition-relevance places (TRPs) and are free-standing, they usually do not constitute the beginning of a larger turn (Gardner 2001, Couper-Kuhlen and Selting 2018). Previous studies on this type of action in Mandarin Chinese have found that these utterances may also be produced to invite more talk and to format displays of information receipt and listener status (Oralova 2016, Gao 2007, Zheng 2007).

A closer look at the 318 instances of the use of monophthongic and nasal particles that format continuers in the corpus utterances shows that these utterances commonly feature a constant or falling pitch contour, and that [ɐ] is the most frequently used phonetic format (Figure 3). Notably, [a], [ɔ] and [m] as well as other formats are also used to format the action.

3.2 Positive response tokens

Positive response tokens (also affirmative tokens) are utterances that, in contrast to continuers, not only display information receipt and listener status, but also constitute a display of affirmation or agreement Couper-Kuhlen and Selting (2018). In Cantonese, monophthongic utterances such as *aa1* (啊) or *ng2* (嗯) can constitute both continuers and affirmative tokens, i.e. these utterances can format interactionally complete affirmative responses.

Figure 4 shows the 41 instances of this use in the data set. Positive response tokens are commonly formatted with a constant or falling pitch contour, using [œ], [ɐ] and [a]. In contrast to continuers, the use of nasal utterances was not observed.

3.3 Change-of-state tokens

Change-of-state tokens format displays that the speaker has moved from a position of unknowing

Figure 3: Continuers: smoothed pitch contour and IPA production format; total n=318

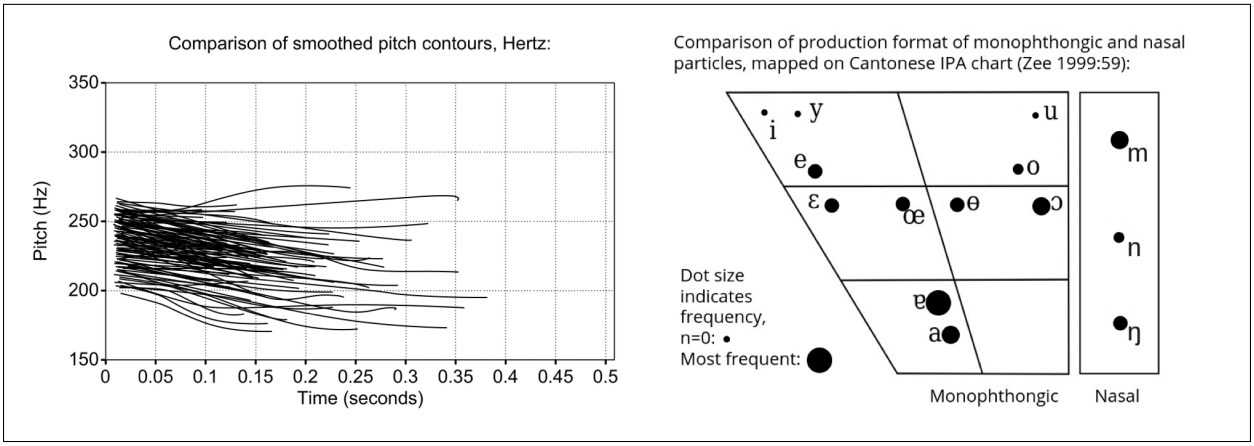
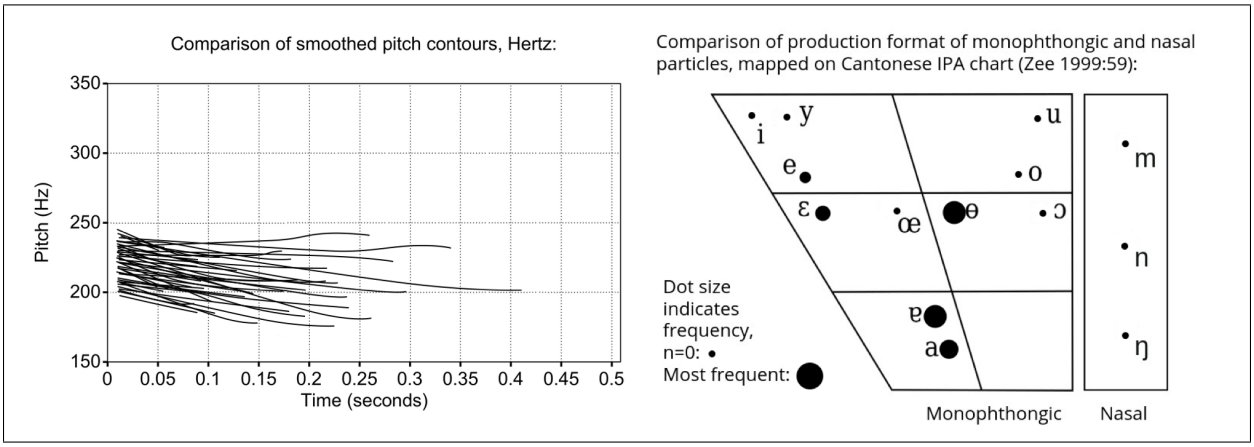


Figure 4: Positive response tokens: smoothed pitch contour and IPA production format; total n=41



to a claimed state of knowing, i.e. they format actions that display understanding or insight (Heritage 1984, 2012). Figure 5 shows 30 instances of the use of monophthonic utterances to format this action. Change-of-state tokens commonly feature a rising pitch contour, and [œ], [o] and [u] are used most frequently in the data set.

3.4 Turn management tokens

Turn management tokens (also turn uptake or turn stalling tokens) format displays of hesitation, reluctance or word search. In the data set (n=68) these utterances commonly feature a constant or falling pitch contour and are formatted using a range of phonetic formats [ɐ],[œ],[θ],[o],[u] and [ɔ]. Notably, neither pitch nor phonetic format appears to be a distinctive feature here, indicating that these actions may be routinely formed by relying on other (possibly sequential) properties.

3.5 Repair initiators

Repair initiators (also trouble tokens or troublesome hearing tokens) are utterances that format displays of a troublesome hearing, doubt or surprise. In the data set (n=27) these actions are commonly formatted featuring a rising pitch contour using [a] or [ɐ].

4 Discussion

Five discourse-interactional functions that can be formatted using minimal (monophthongic and nasal) utterances in turn-initial position have been identified. Based on the analysis of sequential position and

Figure 5: Change-of-state tokens: smoothed pitch contour and IPA production format; total n=30

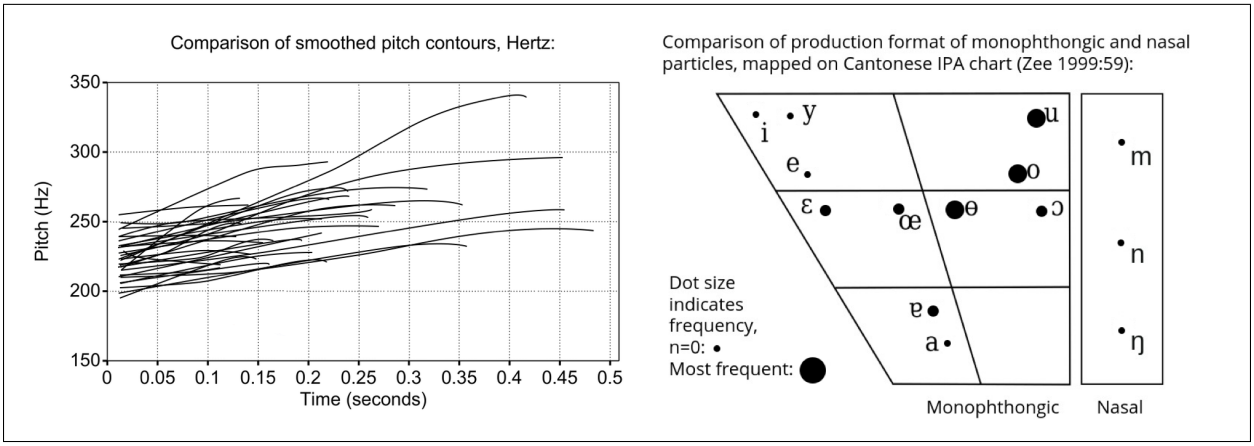


Figure 6: Turn management tokens: smoothed pitch contour and IPA production format; total n=68

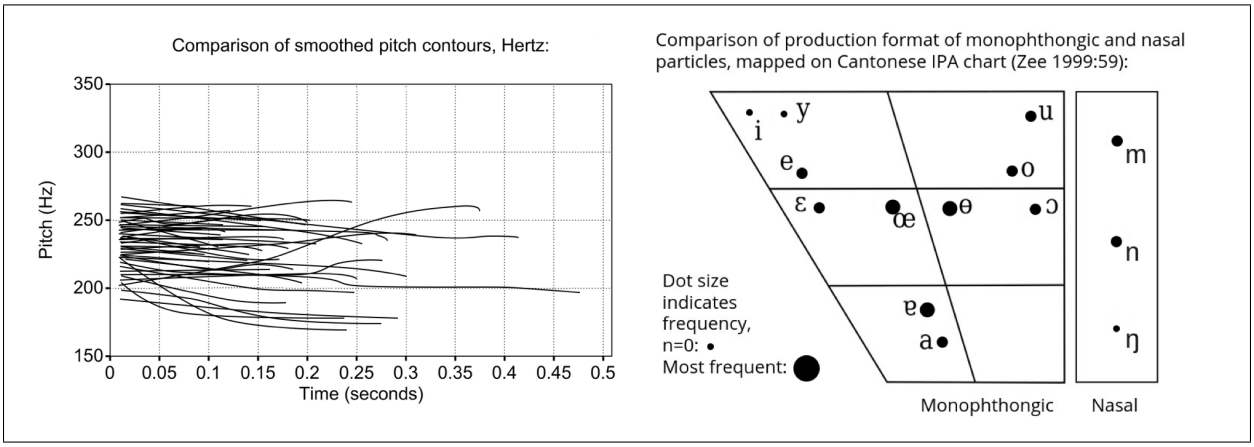
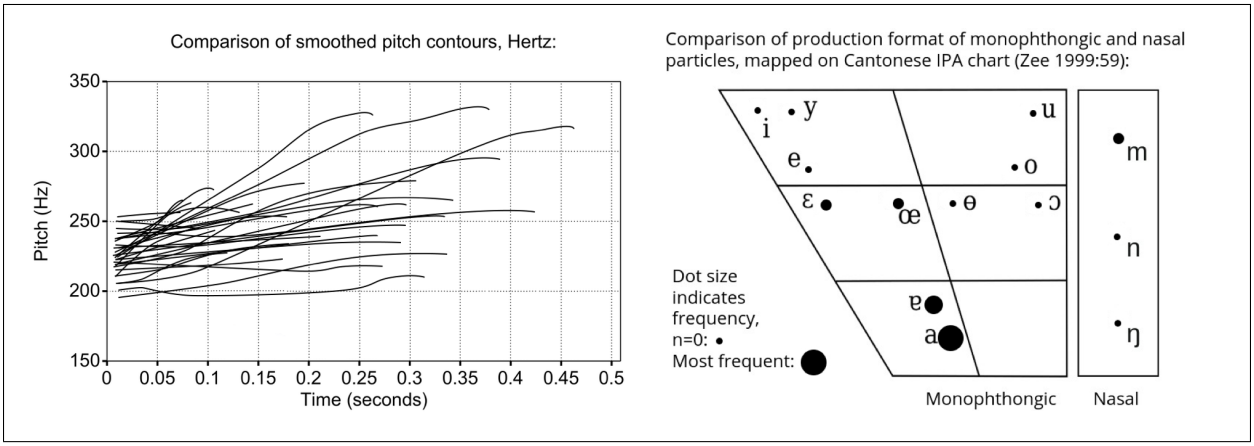


Figure 7: Repair initiators: smoothed pitch contour and IPA production format; total n=27



production format of these utterances in 20 hours of corpus data, I show that the participants format each function using a range of prosodic-phonetic formats that varies in scope, some more constrained than others. I provide an overview of the functions that these utterances can constitute that draws on previous work in interactional linguistics (e.g. Couper-Kuhlen and Selting 2018) and that, in contrast to existing annotation guidelines, is directly grounded in naturally-occurring data. I present an overview of a discourse-functional annotation of 484 usage instances and, focusing on pitch contour and phonetic format, examine the in-situ relationship between action formation and utterance format (Levinson, 2013):

(1) Continuers are produced using a relatively wide range of pitch and phonetic formats, ranging from falling to constant pitch contour and including a range of phonetic formats, with [ɐ] being the most frequent. Continuers are also the most frequent utterance in the data set, making up around 65% of all annotated particles.

(2) Positive response tokens also exhibit constant and falling pitch contours but are more constrained to fewer phonetic formats ([ɐ],[a] and [θ]).

(3) Change-of-state tokens appear to commonly feature a rising pitch contour and are relatively constrained to [θ], [o] and [u].

(4) Turn management tokens feature constant and falling pitch contours but are less constrained in terms of phonetic format, they cover a relatively large range of vocalic and nasal utterance formats.

(5) Repair initiators commonly feature a rising pitch contour and are relatively constrained to [a] and [ɐ].

5 Conclusion

Based on the analysis of recordings of real-world everyday talk-in-interaction, I show that minimal turn-initial utterances of a similar production format can constitute different discourse-interactional functions in different usage environments - a crucial limitation of annotation approaches that rely on direct form-to-function mapping. I conclude that, in order to annotate discourse-interactional functions of minimal turn-initial utterances with reasonable accuracy, at least three structural dimensions have to be taken into account: sequential position, pitch contour and (phonetic) production format. If only pitch contour and production format are considered, good results can be achieved for some functions that appear to be more constrained in their format (such as change-of-state tokens and repair initiators). Other functions, however, appear to not feature strong distinctive prosodic-phonetic properties, which requires their sequential position to be taken into account in order to accurately annotate their respective function (such as turn-management tokens and continuers). The data set shows that participants produce different functions by jointly relying on a range of structural dimensions that are (at least) both sequential and prosodic-phonetic in nature, and that differ in scope for each discourse-interactional function. I hope that this preliminary study provides a useful starting point for further explorations of minimal particles and their involvement in the intricate processes of formation and ascription that participants routinely rely on in natural conversation.

References

Chao, Y. R. (1965). *A grammar of spoken Chinese*. University of California Press.

Cheung, S. H. N. (2007). *A Grammar of Cantonese Spoken in Hong Kong*. Chinese University Press.

Couper-Kuhlen, E. and Selting, M. (2018). *Interactional linguistics*. Cambridge University Press.

Gao, L. (2007). Cong gongneng yuyanxue jiaodu tanxi “en” de shengdiao moshi yu qibiao yi gongneng zhi guanlian (Exploring the tone pattern of ‘en’ and the relationship with its ideographic function from functional linguistic perspectives). *Anhui Literature*, (10):158–159.

Gardner, R. (2001). *When listeners talk: Response tokens and listener stance*, volume 92. John Benjamins Publishing.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. *Structures of Social Action: Studies in Conversation Analysis*, pages 299–345.

Heritage, J. (2012). The epistemic engine: Sequence organization and territories of knowl-

edge. *Research on Language & Social Interaction*, 45(1):30–52.

Hu, M. (1987). *An Exploration of Beijing Mandarin (*北京话初探*)*. The Commercial Press (商务印书馆).

Killingley, S.-Y. (1993). *Cantonese*, volume 6. Lincom Europa.

Lee, J. W., Tao, H., and Lu, P. (2017). Transcribing Mandarin Chinese Conversation: Linguistic and Prosodic Issues. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology (*예술인문사회융합멀티미디어논문지*)*, 7:787–799.

Levinson, S. C. (2013). Action formation and ascription. In *The handbook of conversation analysis*, pages 103–130. Wiley-Blackwell.

Li, C. N. and Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. Univ of California Press.

Liesenfeld, A. (2018). MYCanCor: A Video Corpus of spoken Malaysian Cantonese. In *11th Edition of the Language Resources and Evaluation Conference (LREC), 7-12 May 2018, Miyazaki, Japan*.

Matthews, S. and Yip, V. (2013). *Cantonese: A comprehensive grammar*. Routledge.

Oralova, G. (2016). *Minimal Response Token en in Mandarin Conversation*. PhD thesis, University of Alberta.

Ping, L., Won, L. J., and Hongyin, T. (2014). Discourse Properties of Some Special Sound Elements and Their Transcription Treatment (现代汉语口语中特殊话语语音成分的转写研究). *Linguistic Sciences (*语言科学*)*, 13(2):113–130.

Tseng, S.-C. (2013). Lexical coverage in Taiwan Mandarin conversation. *International Journal of Computational Linguistics & Chinese Language Processing, Volume 18, Number 1, March 2013*, 18(1).

Wu, R.-J. R. (2006). Initiating repair and beyond: The use of two repeat-formatted repair initiations in Mandarin conversation. *Discourse Processes*, 41(1):67–109.

Xu, J. (2009). *Displaying overt recipiency: Reactive tokens in Mandarin task-oriented conversation*. PhD thesis, University of Nottingham.

Xudong, D. (2008). The use of listener responses in Mandarin Chinese and Australian English conversations. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 18(2):303–328.

Zee, E. (1991). Chinese (Hong Kong Cantonese). *Journal of the International Phonetic Association*, 21(1):46–48.

Zheng, Y. (2007). "En" de huayu gongneng fenxi (An analysis of the discourse function of "en"). *Nanfang lunkan*, (10):56–57.