Study on “Ng, A” Type of Discourse Markers in Standard Chinese

Zhigang Yin, Aijun Li, Ziyu Xiong

Phonetics Lab, Institute of Linguistics, Chinese Academy of Social Sciences
yinzgh@cass.org.cn, liaj@cass.org.cn, xiongzy@cass.org.cn

Abstract

This paper attempts to give a detailed study on discourse markers of the “.imag/ı” type in Chinese spontaneous dialogues. Discourse markers (DMs) are considered to have coherent function to the context. It is treated as a kind of linguistic modality to express interactive information or a kind of nonlinguistic modality such as head and body gestures. The discourse markers have aroused many discussions from different aspects of grammar, semantics, cognition and pragmatics. However, this paper will focus on those DMs like “imag” and “ı” from multi-functional and hierarchical points of view and investigate their phonetic relatives or patterns.

Index Terms: Discourse markers, “imag & ı”, pragmatic function, turn-taking, emotion, prosody.

1. Introduction

Since the 1970s, studies on discourse markers (DM) have drawn great attention. According to Schiffrin’s definition, discourse markers are “sequentially dependent elements which bracket units of talk”[2].

Many languages have their “typical” discourse markers, for example, “you know, you see, well” in English and “然后，那么” in Chinese. However, a kind of DM is used universally in all languages. They are some “primary interjections” just like “um”, “oh”, “aha”, “mhm”, “ı”, etc. In spontaneous Mandarin speech, “ı (ı)” and “ı(ı)” are the two most frequently used interjections, so we use the term “imag & ı” type of discourse markers to stand for this kind of DM.

In this paper, “imag & ı” type of DMs will be researched systematically from the aspects of pragmatic function, turn-taking, emotional psychology and phonetics.

Because the “imag & ı” type of DMs mainly appear in spontaneous speech, we used a 3 hour conversation corpus called CADCC which was produced by the Phonetics Laboratory, Institute of Linguistics, Chinese Academy of Social Sciences. Part of the data was selected including 40,000 syllables, for 3 pairs of male interlocutors and 3 pairs of female interlocutors. All the speech data were phonetically and linguistically annotated.

Statistical analysis was done on the annotated data, which showed that the DMs of “imag & ı” type appear 689 times, accounting for 1.7% of the total syllables in the corpus. The pronunciation distribution of “imag & ı” type DMs is shown in table 1.

Table 1. Pronunciation variations of “imag, ı” type of DMs

<table>
<thead>
<tr>
<th>IPA</th>
<th>Ratio</th>
<th>IPA</th>
<th>Ratio</th>
<th>IPA</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ñ</td>
<td>19.91%</td>
<td>m</td>
<td>1.89%</td>
<td>you</td>
<td>0.29%</td>
</tr>
<tr>
<td>c</td>
<td>18.60%</td>
<td>ei</td>
<td>1.16%</td>
<td>an</td>
<td>0.15%</td>
</tr>
<tr>
<td>a</td>
<td>15.99%</td>
<td>hai</td>
<td>1.16%</td>
<td>i</td>
<td>0.15%</td>
</tr>
</tbody>
</table>

2. Pragmatic function

2.1. Classification and statistic analysis

Based on Xu’s research[3], the pragmatic function of “imag & ı” type of DMs is divided into 18 categories shown in table 2.

Table 2. Classification of functions of “imag, ı” type of DMs

<table>
<thead>
<tr>
<th>Definition</th>
<th>Symbol &amp; Occurrence rate</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thinking over wording</td>
<td>THIN 25.5%</td>
<td>Sound produced in brief pondering over wording at the moment of turn-taking or during conversion when the right words cannot be found or when the speaker is not sure about his/her wording.</td>
</tr>
<tr>
<td>Unconscious Feedback (Back channel)</td>
<td>Feed 18.1%</td>
<td>B’s feedback to A’s uttering “process” (not necessarily to A’s uttered “content”).</td>
</tr>
<tr>
<td>Simple confirming</td>
<td>Conf 18.7%</td>
<td>B’s confirming on the content of A’s uttering or questioning.</td>
</tr>
<tr>
<td>Exclamation</td>
<td>Excl 10.9%</td>
<td>Relatively obvious emotional mood.</td>
</tr>
<tr>
<td>Emphasis (or to attract attention)</td>
<td>Emph 6.7%</td>
<td>With prominent phonetic features, which is used to greet, attract attention or enhance expressive capacity of utterance.</td>
</tr>
<tr>
<td>Taking a tumble</td>
<td>Tumb 6.5%</td>
<td>Understand or know all of a sudden, similar to brief confirming but with intensified mood and the feeling of taking a tumble.</td>
</tr>
<tr>
<td>Initiating topics</td>
<td>Init 6.1%</td>
<td>Similar to thinking over wording but contributing more to new turn and helping to clue the other party that the speaker is about to talk.</td>
</tr>
<tr>
<td>Surprised</td>
<td>Surp 1.6%</td>
<td>Sound of surprise produced in intensified emotion, displaying unexpectedness to the other party’s utterance or question.</td>
</tr>
<tr>
<td>Pending</td>
<td>Pend 1.5%</td>
<td>Sound produced when, for example, the speaker is bored as in mutual silence.</td>
</tr>
<tr>
<td>Sudden bethinking strangeness</td>
<td>Sudd 1.3%</td>
<td>The speaker thinks of sth. suddenly.</td>
</tr>
</tbody>
</table>
We found in the statistical analysis that there is a certain correlation between the pragmatic functions of "ng, a" type of DMs and their pronunciations, for example, the function of "thinking over wording (THIN)" is mostly achieved through the "uh-like" syllable "e"(with an occurrence rate of 59%), followed by "ah-like" syllables (ng - 14%, en - 13%) as well as "ea (E)-like" syllables, "EA" (6%); to express the function of "simple confirming (CONF)" the "ah-like" syllable "a" (44%) is most commonly used, followed by the "Oh-like" syllable “ao” (25%) and the “ah-like” syllable “ng” (18%).

2.2. Relationship between Pragmatic function and the pronunciation

We found in the statistical analysis that there is a certain correlation between the pragmatic functions of “ng, a” type of DMs and their pronunciations, for example, the function of “thinking over wording (THIN)” is mostly achieved through the “uh-like” syllable “e” (with an occurrence rate of 59%), followed by “ah-like” syllables (ng - 14%, en - 13%) as well as “ea (E)-like” syllables, “EA” (6%); to express the function of “simple confirming (CONF)” the “ah-like” syllable “a” (44%) is most commonly used, followed by the “Oh-like” syllable “ao” (25%) and the “ah-like” syllable “ng” (18%).

3. Roles in turn-taking system

The concept of turn is first advanced by Sacks et al [4]. It is the basic structural unit of conversation, consisting of word, phrase, sentence or greater language units. The turn-taking system of Sacks includes a turn-constructional component, turn-allocation component and turn-taking rules.

We studied the roles played by DMs of the “ng & a” type in turn-taking system; 8 categories of discourse-turn functions are identified and their occurrence rates are statistically analyzed as shown in table 3.

Table 3. Symbols for turn-taking function

<table>
<thead>
<tr>
<th>Definition</th>
<th>Symbol &amp; Occurrence rate</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiating or taking over turn</td>
<td>INIT 20.8%</td>
<td>1. The speaker initiates a new discourse-turn structure from silence; 2. One speaker ends discourse turn or transfer turn to the listener, who takes the turn to talk.</td>
</tr>
<tr>
<td>Holding turn</td>
<td>HOLD 32.7%</td>
<td>When, Out of incoherence in thinking or other reasons, the speaker cannot express himself continuously but does not want to end his turn, he will take</td>
</tr>
</tbody>
</table>

Table 4. Turn-taking function vs. pragmatic function

<table>
<thead>
<tr>
<th>Turn symbol</th>
<th>function</th>
<th>corresponding pragmatic functions (occurrence rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INIT</td>
<td>Initiating or taking over turn</td>
<td>INIT(28%), EXCL(27%), …</td>
</tr>
<tr>
<td>HOLD</td>
<td>Holding turn</td>
<td>THIN(66%), …</td>
</tr>
<tr>
<td>STOP</td>
<td>Stopping turn</td>
<td>SECO(56%), …</td>
</tr>
<tr>
<td>FEED</td>
<td>Feedback utterance</td>
<td>&quot;FEED&quot; (43%), &quot;CONF&quot; (40%), …</td>
</tr>
<tr>
<td>GRAB</td>
<td>Grabbing turn</td>
<td>TUMB(44%), …</td>
</tr>
<tr>
<td>CHAN</td>
<td>Topic changing</td>
<td>SUDD(99%), …</td>
</tr>
</tbody>
</table>

We also find a correlation between turn-taking functions and the DMs’ pronunciations. The feedback (FEED) is mainly realized as /ng, a/ and /ao/; holding turn is primarily uttered in /e/ and /EA (ê), en, ai, "ai, e, EA (ê), a/ are employed to display initiating and taking over a turn.

4. Emotional and psychological analysis

4.1. Three subordinate psychological layers

People’s behaviors in dialogue can be roughly divided into two parts: listening stage (understanding and estimating the utterance of the other party simultaneously, and forming trusting-doubting judgment) and speaking stage (expressing one’s own opinion, showing accepting-rejecting communication orientation). Apart from these two stages, linguistic behaviors of the interlocutors are also influenced by emotional status.

The three subordinate psychological layers are defined as emotion layer, trusting-doubting layer, and acceptance - rejection layer, which relate to the emotional status, understanding-judgment behavior, and speech communication behavior respectively.

(A) Emotional layer: exploring the impact of the speaker's emotional polarity (positive, neutral, and negative emotion, represented respectively by ‘+’, ‘0’ and ‘−’ ) and emotional intensity (high, medium and low represented by ‘0’, ‘1’, ‘2’).
(B) Trusting-doubting layer: exploring the impact of the speakers' believing-doubting about their spoken content on the use of "ng & a" DMs (divided into 5 degrees of intensity, "believe in deeply", "believe", "undecided", "hesitating or doubting " and " do not believe", represented respectively by 2, -1, 0, +1, +2).

(C) Acceptance-rejection layer: exploring the relationship between the use of "ng & a" type DMs and the speakers' acceptance-rejection psychology (divided into 5 levels of mildness, "approving";" accepting", "undecided", "hesitating" and "rejecting", represented respectively by 2, -1,0, +1, +2).

4.2. Analysis according to the three subordinate psychological layers

Neutral emotions take up 91% when viewed from the emotion polarity; and low activation (marker is 0) accounts for 69.1% when viewed from the degree of emotion activation.

Occurrence times of DMs in the trusting-doubting layer fall mainly in normal distribution: +2 (believing in deeply, 1.7%), +1 (believe, 18.4%), 0 (undecided, 54.6%), 1 (hesitating and doubting, 24.5%); 2 (do not believe, 0.4%).

Occurrence times of DMs in the acceptance-rejection layer accord basically with the normal distribution: +2 (approving, 1.9%), +1 (accepting, 25.4%), 0 (undecided, 57.9%), -1 (hesitating, 13.9%), -2 (rejecting, 0.9%).

There is correlation between pragmatic function and psychological layers. In each subordinate psychological layer, the categories of pragmatic function can be sorted according to their psychological arousal degree. For an instance, in the trusting-doubting tier, several pragmatic functions with high occurrence frequency can be sorted from a strong trusting degree to a weak degree as: "taking a tumble" > "simple confirmation" > "feedback to acknowledgement" > "Initiating Discourse" > "thinking over wording" > "strangeness" > "unexpectedness (surprised)."

Relations among the three subordinate psychological layers are also observed.

(1) Relationship between the trusting-doubting layer and the emotional layer

Polarities (positive, neutral, negative) of the trusting-doubting layer and the emotional layer are not correlated (r = 0.109, P = 0.01); The relationship between the Cognition-trusting-doubting layer and the activation degree of the emotional layer (high, medium, low) is similar (r = 0.112, P= 0.01).

(2) Correlation between the Acceptance-rejection layer and the Emotional layer:

Polarities (positive, neutral, negative) of the acceptance-rejection layer and the emotional layer are not significantly correlated (P= 0.01, r= 0.249); the relationship between the acceptance-rejection layer and activation degrees of the emotional layer is also similar (P= 0.01, r= 0.245).

(3) Correlation between the Acceptance-rejection layer and the Cognition-trusting-doubting layer

Significant positive correlation exists between the Acceptance-rejection layer and the Cognition-trusting-doubting layer (p= 0.01, r = 0.68). A psychological explanation of this positive correlation is that when people believe in what others said, they tend to accept others, and when they suspect others, they tend to refuse and vice versa.

5. Segmental and prosodic analysis

5.1. Pronunciation

We analyzed the relationship between the “ng & a” type DMs' pragmatic function, emotional function and their pronunciation, getting the following results:

(1) The “ng & a” type DMs with stronger pragmatic function and greater information capacity will be pronounced more loudly (greater opening degree) and have more complex pronunciation forms (containing more phones).

(2) The “ng & a” type DMs with higher activation or arousal degree of emotion will be pronounced more loudly (greater opening degree) and have more complicated forms of pronunciation (containing more phones).

5.2. Duration

Since there are differences among different speakers' speaking rate and average syllable length, we use Z-SCORE to normalize the speaker's syllabic duration:

\[ T_x = \left( X_x - \mu_x \right) / \sigma_x \]

\( T_x \) refers to the original duration of a certain syllable of speaker X; \( \mu_x \) refers to the average duration of all syllables of speaker X, \( \sigma_x \) is the standard deviation of all syllabic duration of speaker X; \( T_x \) refers to the normalized duration of the syllable.

Inspecting the average duration of "ng & a" DMs, we find that the average duration of "ng & a" DMs' is longer than that of the overall syllables (Normalized duration is 1.196, and the reference standard duration is 0).

Through multi-variance analysis (ANOVA), we found that pragmatic function, turn-taking, polarity of emotion, trusting-doubting information and acceptance-rejection information all have significant impact on duration(P<0.05), while emotional activation does not (P> 0.05).

In addition, we investigated the impact of various pragmatic functional categories on DMs’ durations (see figure 1). We can see that only durations of ‘Sudd’ and ‘Repa’ are shorter than the average, while others are all longer than the average duration.

![Figure 1: Duration distribution relating to pragmatic functional categories](image)

5.3. Pitch

In order to eliminate gender and individual differences among speakers' pitch, we normalized the \( F_0 \) data as well, using Shi Feng's [5] 5-tone letter scale:

\[ T = 5 \times (\log_{10} F_0 - \log_{10} \text{min}) / (\log_{10} \text{max} - \log_{10} \text{min}) \]

Where, \( F_0 \) is the fundamental frequency of the observed phonetic units, ‘\( \text{Fmax} \)’ and ‘\( \text{Fmin} \)’ are the maximum and minimum value of speaker's pitch range respectively.

Similar to duration studies, we ran ANOVAs on all level factors and the pitch.
(1) Pragmatic functions have no significant impact on the mean value of pitch, but have significant impact on the pitch range, upper- and lower- limits.

(2) Turn-taking functions have no effect on the pitch mean, range, upper and lower limits.

(3) Emotional activation has significant impact on the pitch mean, upper and lower limits, but not on pitch range.

(4) Acceptance-doubting has significant influence on the pitch mean, but not on the other pitch cues.

(5) Acceptance-rejection has significant impact on the pitch mean, upper and lower limits as well as pitch range.

Moreover, we studied the relations between the pragmatic functions and pitch contour and found the following results:

(1) Falling tones are usually adopted in expressing the function of confirming or acceptance, which include "simple confirming", "back channel", "exclamation" and "taking a tumble", etc.

(2) Level tones are adopted in non-expressive functions, such as "thinking over wording," "initiating topics," etc.

(3) Rising tones are used in arousing attention or displaying interrogation and puzzling, such as part of "emphasis", "thinking of suddenly," "strangeness" and "checking for interrogation and puzzling, such as part of "emphasis", "thinking of suddenly," "strangeness" and "checking for "emphasis", "thinking of suddenly," "strangeness" and "checking for

We analyzed F0 contour in pragmatic functions and obtained 14 typical pitch patterns of all 18 pragmatic categories (Fig. 2).

Figure 2: Pitch patterns relating to pragmatic functions (x-axis: Normalized duration)

6. Perceptual Experiment and Conclusion

36 utterances with different "ng, a" type DMs and pragmatic functions were first selected, and then all the DMs were substituted by the synthesized ones according to the above prosodic patterns relating to pragmatic functions (Praat was used as the synthesis software).

6 subjects judged the expressive difference between the original and synthesized utterances, showing 92% have no difference.

Another perceptual experiment was conducted to verify the impact of the context and prosodic patterns on pragmatic functions. 5 utterances bearing DM /ng/ were selected with the 5 functions of “Confirmation, Feedback, Exclamation, Repair and Surprise”. Another /ng/ with the “thinking” function in average prosodic state was chosen to replace the former 5 DMs. Perceptual results show that 50% of the utterances changed their functions, in which 80% were perceived to express ‘thinking’. This demonstrates that prosodic features contribute to the perception of pragmatic function. Carefully checking those 50% unchanged utterances, we found their DMs have rather similar pitch patterns to that of “thinking”(level or slightly falling).

So what we may conclude from the perceptual experiment is that (1) Both context and prosodic patterns have impact on pragmatic functions; (2) Slight change of DM’s prosodic pattern can not change its pragmatic function; great change of the prosodic pattern will cause its function to change while ignoring the context. (3) Prosodic pattern of DM should accord with its context.

We suggest that the results would be useful in human-machine interaction systems. This research is supported by the National 863 project #2006AA01Z138.

7. References


