Emphatic Stress as Epistemic Conflict: A case study of Mandarin Chinese

Yurie Hara\textsuperscript{1}, Shigeto Kawahara\textsuperscript{2}, and Yuli Feng\textsuperscript{1}

\textsuperscript{1} City University of Hong Kong
\textsuperscript{2} Rutgers University
y.hara@cityu.edu.hk kawahara@rci.rutgers.edu yuli.feng@student.cityu.edu.hk

Abstract. This paper examines the semantics of sentence-final stress in Mandarin Chinese and gives a uniform account of the distribution and interpretation of emphatic stress. A phonetic production study shows that the use of sentence-final stress by Mandarin speakers is sensitive to the evidentiality of the context. A semantic rating study shows that listeners judge sentences with evidenced contexts as more natural when they are accompanied by sentence-final stress. Based on these results, we argue that a sentence-final stress gives rise to an alternative proposition which is obtained by switching the polarity of the sentence, and that the sentence-final stress indicates an epistemic conflict between the speaker’s belief and the incoming proposition at issue.

1 Prosodic Stress in Mandarin Chinese

In Mandarin Chinese, prosodic stress on a constituent can indicate an emphatic focus (c.f., Liu and Xu [2005], Peng et al. [2006]). In (1), the subject NP \textit{Weili} receives emphatic stress (marked as \textit{es}), which gives rise to a narrow focus on \textit{Weili}; that is, (1) indicates that among alternative individuals, it is Weili that sells bacon, not other people. Phonetically, the emphatic stress is realized via local expansion of the pitch range of focused syllables as well as post-focus compression, as depicted in Figure 1 (taken from Peng et al. [2006]). Semantically, emphatic stress gives rise to alternative propositions. In the case of constituent stress like (1), the alternative propositions are obtained by replacing the focused constituent with alternative constituents of the same semantic type (e.g., type e in (1), see Rooth [1985]). For instance, we obtain the following alternative set for (1): Alt([[\textit{Weili} sells bacon]]) = [[\textit{Weili sells bacon}], [Zhansang sells bacon], [Xiaomei sells bacon], ...]. Further pragmatic strengthening will yield exhaustification [see Schulz and van Rooij, 2006].

\begin{equation}
[\text{es}\, \text{Weili}\, \text{mai}\, \text{la}\, \text{rou}] \, \text{sell}\, \text{bacon}.
\end{equation}

\textit{‘WEILI sells bacon’ (It is Weili, not others, that sells bacon.) (adapted from Peng et al. [2006])}
(a) Narrow focus on WeiLi
(b) “Out of the blue” statement

Fig. 1. Pitch expansion due to constituent stress in Mandarin Chinese [Peng et al, 2006; p. 52]
The same pitch expansion can occur sentence-finally. In (2), stress (marked as sfs) falls on the last syllable of a falling positive declarative, and in (3), a rising negative declarative. This paper investigates the phonetics and semantics of these types of sentences with sentence-final stress. Semantically, introspection-based data suggests that when the verb-final syllable bu51 is pronounced with stress, the utterance indicates that the speaker had a previous belief toward a proposition with the opposite polarity. More specifically, final stress in a falling declarative like (2) gives an impression of the speaker’s surprise, while final stress in a rising declarative like (3) indicates the speaker’s incredulity about the embedded proposition.

(2)  Ni21 ping35 shi35 pao21 [sfs bu51]!
You usually   jogging
‘You usually jog!’ (Falling Positive Declarative)
(I believed that you don’t usually jog.)

(3)  Ni21 ping35 shi35 bu51 pao21 [sfs bu51]?
You usually   NEG jogging
‘You don’t usually jog?’ (Rising Negative Declarative)
(I believed that you usually jog.)

Building on the discussion of the focus effect of constituent stress observed by Peng et al. [2006], we propose that the emphatic stress is licensed only if the alternatives are salient in the context. In the case of sentence-final stress, stress indicates a focus-marking on the polarity of the sentence; hence the alternative proposition is obtained by switching the polarity of the uttered sentence. The alternative proposition is salient because it is part of the speaker’s previous belief. In sum, constituent stress and sentence-final stress can both be characterized as triggering salient alternatives.

When does the speaker have a belief toward a particular proposition? The most plausible situation is when the speaker has some reason to believe the proposition, i.e., evidence which supports the proposition. That is, the strength of evidence determines the speaker’s epistemic states and therefore influences the change of belief. Therefore, our proposal predicts that a sentence uttered with sentence-final stress is more natural when the speaker has evidence for the alternative proposition and thus believes it. Our phonetic production study shows that evidential intonation indeed exhibits characteristics of focal stress in Mandarin Chinese. The semantic naturalness rating study also shows that Mandarin listeners judge the sentences with evidence more natural when they are accompanied by sentence-final stress. We thus conclude that sentence-final stress indicates an epistemic conflict between the proposition at issue and the speaker’s previous belief.

2 A production study

We started with a production study to investigate whether Mandarin Chinese speakers use sentence-final stress to reflect the evidentiality of context. Liu and
Xu [2005] show that focal stress in Mandarin Chinese is phonetically characterized by an elevated and expanded pitch range. The current production study aimed to investigate whether evidentiality would induce a similar type of focal stress.

2.1 Method

Ten Chinese speakers were instructed to read 10 interrogative sentences in Mandarin Chinese. The target sentences mainly consisted of words with sonorants to obtain unperturbed F0 contours. By attaching the question particle *ma* to the final position, the target sentences are unambiguously of the interrogative type. The particle *ma* has the neutral (i.e. lexically underspecified) tone. The penultimate word of every target sentence has the falling tone, as exemplified in (5). The speakers were told to read the target sentences in two conditions; one where there was evidence for the positive answer (the evidenced condition) and one where there was no such evidence (the neutral condition). A context which differentiates the evidentiality status was provided for each target sentence, as illustrated in (4-a) and (4-b). Their speech was recorded using a digital recorder (Tascam HD-R1 solid state recorder) and a microphone (Shure dynamic microphone SM48) at 44.1k sampling frequency.

(4) a. Evidenced Context: A invites B to a folk gig, yet B hesitates as if she has no interest. However, A often see B’s posts on a folk music forum and knows B collects folk music albums, so A asks B,...
   b. Neutral Context: A invites B to a folk gig, yet B hesitates as if she has no interest. And A asks B,...

(5) Nǐ₂¹ bù₅¹ tīng₅₅ yǎo₅₅ yīn₅₅ yuè₅₁ ma?
   ‘You don’t listen to folk music?’

To analyze the intonation contours, the recorded sentences were first annotated using Praat (Boersma and Weenink [1992-2011]). The difference between the two conditions was most apparent in the penultimate word and the final word (i.e., *[ma]*) so our analysis focused on these two last intervals. In order to calculate the F0 movements, these two intervals were each divided into ten equally-spaced windows, and the average F0 value was calculated within each window. A linear mixed model was used to assess the difference between the two conditions. Speakers and items were added as random factors in the analyses.

---

3 Although all the participants reported that they were native speakers of Mandarin Chinese, one of the speakers (originally coded as Speaker 1) had a non-Mandarin accent, and the data of this speaker was thus excluded from the following analysis.

4 Our thanks to Yizhou Lan, Sunyoung Oh, Melanie Pangilinan and Shuangdi Zhang for their assistance on this process.

5 We also calculated the intensity (average and maximum) and duration of [ma] (cf. Liu and Xu [2005]), which revealed no differences between the two conditions. We did not calculate these values for the penultimate intervals, because the onset boundaries
2.2 Results

Figure 2 illustrates the schematic F0 contours of the penultimate and final words in both conditions (the evidenced condition and the neutral condition). The first ten windows track the pitch contours of the penultimate words, and the last ten windows, those of the final word (i.e., [ma]). We observe that the pitch contours in the evidenced condition show generally elevated and raised pitch ranges. For the majority of speakers, the F0 contours are generally higher and show more pronounced movements in the evidenced condition than in the neutral condition. Speakers 3 and 8 did not seem to show differences between the two conditions; they may have failed to pay attention to the contexts we provided for each sentence. The data from Speaker 1 is not shown here, because the speaker did not seem to be a native speaker of Mandarin Chinese (see footnote 3).

As a measure of pitch range, we calculated the minimal F0 and maximum F0 of the rise found at the beginning of the penultimate words, because they seem to define the lowest and highest pitch points for all speakers. In Figure 2, for those speakers who show differences between the two conditions, the differences manifest themselves in the rising F0 movement in the penultimate words. We did not base our calculation of the pitch range on the final words, because Speaker 6 shows a reversal in this position.

Figure 3 thus illustrates the average F0 minima and maxima of the F0 values of the rise found at the beginning of the penultimate words in the two conditions, averaging over all the speakers and all the items. Both minima and maxima are raised in the evidenced condition (minima: \( t = -2.082, p < .05 \); maxima: \( t = -3.755, p < .001 \)). Figure 4 illustrates the size of the pitch range (maxima-minima), which shows that the evidenced condition showed a larger pitch range than the neutral condition (\( t = -2.043, p < .05 \)).

2.3 Discussion

The results of the production study indeed show that evidentiality of the context is characterized by the elevated (raising of minima and maxima) as well as expanded pitch range (the larger difference between minima and maxima). These phonetic characterizations of sentence-final stress match well with those identified by Liu and Xu [2005] for focal stress in Mandarin Chinese. The results thus show that the use of sentence-final stress by Mandarin speakers reflects the evidentiality of the context.

---

of these words were hard to draw, as these intervals often started with approximants (e.g., [yue] 'music').
Fig. 2. Schematic pitch contours of the penultimate and final words for each speaker, averaged across 10 items. The first ten windows represent the F0 movements of the penultimate intervals. The last ten windows represent those of the final intervals (i.e. [mə]). The y-axis scales are different for each speaker.
**Fig. 3.** The F0 minima and maxima of the tonal movements at the beginning of the penultimate words. The error bars represent 95% confidence intervals.

**Fig. 4.** A comparison of the pitch ranges between the two conditions, measured by F0 maxima-F0 minima. The error bars represent 95% confidence intervals.
3 A naturalness rating study

Having established that speakers use sentence-final stress to indicate evidentiality, the next experiment investigated whether the two types of intonational patterns—with and without sentence-final stress—would impact the interpretation of sentences in terms of their evidentiality. To address this question, the second experiment used a semantic naturalness rating paradigm.

3.1 Method

Native speakers of Mandarin judged the naturalness of the combination of two factors—stress and evidentiality. Each stimulus consisted of a context, which distinguished evidentiality, as exemplified in (6-a) and (6-b), and a target rising negative declarative like (3) (repeated here as (7)) with or without final stress. Each condition had 16 items. 64 target sentences plus 64 fillers were created. The stimuli were recorded in a sound-attenuated room and produced by a native speaker of Mandarin Chinese. The speaker was specifically instructed to manipulate the presence of stress (not evidentiality). The order of the stimuli was pseudorandomized so that no minimal pair sentences appeared next to each other.

(6) a. Evidenced Context: B says he rarely exercises. A often catches sight of B jogging in the park, so A asks B,...
   b. Neutral Context: B consults A on how to keep fit. And A asks B,...

(7) Ni21 ping25shi35 bu51 pa021bu51?
   You usually NEG jogging
   ‘You don’t usually jog?’ (Rising Negative Declarative)

20 native speakers rated the naturalness of the match of the target sentences against the contexts on a 1-to-5 scale from completely unnatural to completely natural. The participants listened to the stimuli using headphones in a sound-attenuated room. Perception, an assessment management software, was used to present the stimuli. To analyze the results, a general linear mixed model was run, in which stress and evidentiality were the fixed factors and speakers and items were the random factors. If final stress depends on the evidentiality of the context, the dependency is expected to result in a significant interaction between stress and evidentiality.

3.2 Results

Figure 5 shows the average naturalness ratings in each condition. In the evidenced contexts, speakers judged questions with final stress more natural, whereas in the neutral contexts, speakers judged sentences without final stress more natural. Because of this asymmetry, the interaction between stress and evidentiality was significant ($t = 18.70, p < .001$).
3.3 Discussion

The results show that Mandarin listeners prefer to have sentence-final stress for evidenced contexts, while they prefer to not have sentence-final stress in neutral contexts. This asymmetry confirms that sentence-final stress indicates that the speaker has evidence toward the alternative proposition $p$, i.e., the proposition with the polarity opposite to the surface sentence $\neg p$.

Evidentiality and epistemic states are closely related notions where evidentiality encodes the source of information while epistemic states depict the certainty of knowledge [Givón, 1982]. In Gärdenfors’ [1988] probabilistic model of epistemic states, a speaker’s epistemic states consist of propositions arranged in a decreasing order with respect to likelihood, and the likelihood of a proposition $p$ correlates to the evidence the speaker has for or against $p$. Hence, if the speaker has sufficient evidence for $p$, then $p$ will be arranged high in his epistemic states, which will in turn cause belief revision. Therefore, in the evidenced context, it is most natural to assume that the speaker believes $p$, which conflicts with the at-issue proposition denoted by the surface sentence. In short, the sentence-final stress indicates an epistemic conflict because the proposition at issue $\neg p$ is at odds with the speaker’s previous belief $p$.

4 Integrating the results

Both the production study and the rating study show that the use of the sentence-final stress depends on the availability of evidence in the context. The
next questions are: (i) what is the role of stress?; and (ii) how does the interpretation of sentence-final stress interact with the context? In particular, our native speaker consultants report that the sentence-final stress seems to indicate the speaker’s surprise or incredulity. How does this surprise connotation come about? To answer these questions, we first propose that stress in general induces salient alternatives unifying both constituent and sentence-final stress; then we derive the interpretation of the utterances with sentence-final stress in terms of the dynamic update over public beliefs.

4.1 The role of stress


(8) 
\[
\text{[Weili] sells bacon.} \\
\text{(It is Weili that sells bacon)}
\]

(8), when uttered with emphatic stress on the subject \textit{Weili}, yields an exhaustive interpretation, i.e., ‘Only Weili sells bacon’. More specifically, there seems to be a question under discussion, ‘Who sells bacon?’ or a set of alternative propositions, \textit{[Weili sells bacon]}, \textit{[Xiaomei sells bacon]}, \textit{[Xiaohua sells bacon]}. We adapt the framework of Rooth [1992] in deriving the alternatives. In Rooth [1992], stress on a particular constituent indicates a focus marking on the stressed constituent. The existence of the focus in a sentence prompts the focus semantic value \(f\) in addition to the ordinary semantic value \(\cdot\). The focus value can be obtained by abstracting the proposition over the focused element (9-a). However, the focus value cannot directly be the alternative set since the focus value includes all the propositions of the form \textit{sell}(x, bacon), as long as \(x\) is an individual. Rather, the alternative set is constructed from the context and the focus value constrains the form of the members of the alternative set by requiring it to be a subset of the focus value (9-b). In other words, a focus-marking imposes a presupposition on the context so that there be a salient set of alternatives.

(9) a. \(\llbracket [f \text{ Weili} \] \text{ sells bacon.} \rrbracket = \text{sell}(x, \text{bacon}) \times e \in D_e\)

b. Alt(\(\llbracket [f \text{ Weili} \] \text{ sells bacon.} \rrbracket\))
\(= \{\text{sell}(\text{weili, bacon}), \text{sell}(\text{xiaomei, bacon}), \text{sell}(\text{xiaohua, bacon})\}\)
\(\subset \llbracket [f \text{ Weili} \] \text{ sells bacon.} \rrbracket\)

The exhaustive meaning is obtainable by pragmatic strengthening [c.f., Schulz and van Rooij, 2006]. To simplify the picture, let us introduce an exhaustive operator \textit{Exh} which operates over the alternative set and the focused element.

(10) \(\text{Exh(Alt}(p), p) = \forall q=q \in \text{Alt}(p) \land w_\theta \in q \rightarrow q = p\)
For each proposition in the alternative set salient in the discourse, if it is true, it is identical to the surface proposition $p$. Hence, only the surface proposition is true among the alternatives. The crucial point here is that the focus-marking presupposes that there is a salient set of alternatives in the context.

Let us go back to the sentence-final stress, which is the main concern of the current paper. Although sentential-final stress may seem to have a function different from constituent stress, we propose that it is possible to unify the functions of constituent and sentence-final stress. That is, the role of stress is to induce salient alternatives. In case of sentence-final stress, it focus-marks the covert affirmative operator, Ident, which is an identity function of the semantic type $\langle st, st \rangle$. Now, what would be the alternative for Ident? Our answer is the negative operator $\neg$, which reverses the truth value of the proposition. Hence, the alternative set of the sentence $neg-S$ with sentence-final stress is $\{\neg p, \neg \neg p\} = \{\neg p, p\}$, where $[neg-S] = \neg p$.

(11) a. $[\text{You don’t usually jog-}[\neg \text{Ident}]]^f$
   $= \{O(\neg \text{you-usually-jog})| O \in D_{\langle st, st \rangle}\}$

b. Alt($[\text{You don’t usually jog-}[\neg \text{Ident}]]$)
   $= \{\neg \text{you-usually-jog}, \neg \neg \text{you-usually-jog}\}$
   $\subset [\text{You don’t usually jog-}[f]]^f$

Now, as in the case of constituent stress, the existence of the alternative set is a presuppositional requirement imposed by sentence-final stress. The surface proposition $\neg p$ is salient as it is the at-issue proposition, while the salience of the alternative proposition $p$ is dependent on the context. In our two experiments, we have looked at two different contexts, evidenced and neutral contexts. In the evidenced contexts, the speaker has evidence for the alternative proposition $p$, which thus makes $p$ salient, and therefore the use of sentence-final stress is licensed. In contrast, the neutral contexts do not make $p$ salient, and thus the use of sentence-final stress is not licensed as it would not satisfy the presuppositional requirement of sentence-final stress.

4.2 Semantic analysis

To formally model the interpretation of the utterance with sentence-final stress, we employ the framework of Gunlogson [2003], which defines a Stalnakerian [1998] common ground as an intersection of the public beliefs of the conversation participants, i.e., the speaker and the addressee. The common ground is a set of propositions, each of which is a mutual belief of the participants in the discourse. An assertive act is characterized as a proposal to add a new proposition into the common ground [Stalnaker, 1998]. Gunlogson [2003], in analyzing English rising declarative sentences, provides a more fine-grained notion, the public belief, which is relativized to each conversation participant:

(12) With respect to a context $C$,
   a. $p \in \text{PBA}(C)$ iff ‘$A$ believes $p$’ is a mutual belief of $A$ and $B$ in $C$. 

   b. $\text{PBA}(C) = \text{PBA}(D)$ iff ‘$A$ believes $p$’ is a mutual belief of $A$ and $B$ in $C$. 

b. \( p \in PB_B(C) \) iff ‘B believes \( p \)’ is a mutual belief of A and B in C.

[adapted from Gunlogson, 2003]

A linguistic utterance performs an update to this public belief in the sense of Heim [1982]. According to Gunlogson, a rising ‘↑’ declarative like (3) denotes the addressee’s commitment to the embedded proposition, and hence the proposition \( \neg p \) is added in the addressee’s public belief, \( PB_{addr}(C) \) as defined in (13).

\[
\llbracket \uparrow \rrbracket (\neg p)(C) = PB_{addr}(C) + \neg p,
\]

where \( C \) is the current context of the utterance.

As discussed in section 4.1, sentence-final stress focus-marks the affirmative operator of the surface proposition and hence induces a set of alternatives: \( Alt(p) = \{ p, \neg p \} \). The salience of the surface proposition \( \neg p \) is evident as it is uttered by the speaker and committed to by the addressee. Therefore, the definition of sentence-final stress SFS needs to further pin down the salience of the alternative proposition with the opposite polarity \( Alt(\neg p) = \{ \neg p \} \) (\( = \{ p \} \)). In particular, the alternative is salient because it is part of the speaker’s previous belief. Note that the belief here could be public or private; hence we use \( \text{Bel}_{\text{spkr}} \).

\[
\llbracket \text{SFS} \rrbracket (\neg p)(C) \text{ is defined iff } Alt(\neg p) - \{ \neg p \} \subset \text{Bel}_{\text{spkr}}(C_{\text{pre}}),
\]

where \( Alt(\neg p) = \{ \neg p, p \} \) and \( C_{\text{pre}} <_T C <_T C_{\text{post}} \).

In this sense, sentence-final stress is more compatible with the contexts where the speaker has evidence for \( p \). In sum, a declarative sentence uttered with a rising intonation and sentence-final stress has the following interpretation. Final rise indicates that the addressee is committed to the proposition \( \neg p \), which conflicts with the alternative proposition \( p \) which is part of the speaker’s previous belief.

(15) The interpretation of \( \text{NEG-S-SFS-↑} \):

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>( \neg p \in PB_{addr}(C) )</td>
</tr>
<tr>
<td>b.</td>
<td>( { p } \subset \text{Bel}<em>{\text{spkr}}(C</em>{\text{pre}}) ),</td>
</tr>
</tbody>
</table>

where \( \llbracket \text{S} \rrbracket = p \).

Thus, the speaker attempts to resolve the conflict by uttering a rising declarative. In other words, the communicative intent of the speaker goes as follows: I have enough evidence for \( p \); and you are publicly committed to \( \neg p \). Therefore, I request whether you have sufficient evidence to commit yourself to \( \neg p \), i.e., whether you are sure to add \( \neg p \) into the common ground. In contrast, when the speaker has no evidence, then \( p \) does not belong to the speaker’s set of beliefs, i.e., \( p \) is not salient. The speaker can at most make a mere guess about the trueness of the alternative proposition; therefore, there is naturally no epistemic conflict and final stress is less preferred.

Furthermore, our analysis extends to a sentence-final stress in a falling ‘↓’ positive declarative as in (2). A falling declarative denotes the speaker’s commitment to the proposition as defined in (16).

\[
\llbracket \downarrow \rrbracket (p)(C) = PB_{spkr}(C) + p
\]
With the additional emphatic stress, the utterance induces a presupposition as in (17). That is, the speaker had a previous belief that ¬p is the case:

\[(17) \quad [SFS](p)(C) \text{ is defined iff } Alt(p) - \{p\} \subseteq Bel_{spkr}(C_{\text{pre}}),\]
\[
\quad \text{where } Alt(p) = \{¬p, p\} \text{ and } C_{\text{pre}} < T C < T C_{\text{post}}.
\]

The utterance results in the expression of the speaker’s surprise since the newly updated proposition conflicts with the speaker’s previous belief.

\[(18) \quad \text{The interpretation of S-sfs-↓:}\]
\[
\quad a. \quad p \in PB_{spkr}(C) \\
\quad b. \quad \{¬p\} \subseteq Bel_{spkr}(C_{\text{pre}}).
\]

The communicative intent of (16) will then be: I agree to add p into the common ground despite having evidence for ¬p and I am thus surprised that p is the case.\(^6\)

Accordingly, we decompose the meaning of declaratives with sentence-final stress into three components, namely, their propositional content, the final tone, and the stress, where the final tone is the intonational cue of propositional commitment relativized to interlocutors and the stress, the cue of the speaker’s previous belief.

5 Conclusion

An emphatic stress indicates that the alternative propositions are salient. The focus-marking signaled by constituent stress yields a set of alternative constituents with the same semantic type as the focused. Sentence-final stress does not indicate narrow focus on specific constituents; rather, it focus-marks the polarity of a surface proposition and thus makes the proposition with the opposite polarity a salient alternative. In this way we have developed a unified analysis of emphatic stress via focus semantics. One of the causes that renders the alternative proposition salient is when the speaker has evidence which supports the proposition. Two experiments reported in this paper demonstrate the tight relationship between sentence-final stress and the context where the speaker has evidence for the opposing proposition. That is, sentence-final stress expresses an epistemic conflict between the speaker’s previous belief and the proposition at issue. Hence, our study shows how prosody can reflect the speaker’s epistemic states and yield a discourse effect.

\(^6\) The other possible constructions, falling negative and rising positive declaratives, can be analyzed in the same fashion as depicted in (i) and (ii), respectively.

(i) The interpretation of neg-S-sfs-↓:
\[-p \in PB_{spkr}(C) \land \{p\} \subseteq Bel_{spkr}(C_{\text{pre}}).\]

(ii) The interpretation of S-sfs-↑:
\[p \in PB_{addt}(C) \land \neg\{¬p\} \subseteq Bel_{spkr}(C_{\text{pre}}).\]
Bibliography