CHAPTER 3

‘Like Godzilla’

Enactments and formulations in telling a disaster story in Japanese

Alfred Rue Burch and Gabriele Kasper
University of Hawai‘i at Mānoa

The chapter examines how a second language speaker of Japanese tells a disaster story to an L1 Japanese-speaking recipient in ordinary conversation. Drawing on Goodwin’s (2013) notions of lamination and substrates, the study shows how the teller and recipient orient to the story as a stance object by selecting, assembling, and recycling different types of multisemiotic resources, including language forms, cultural references, prosody, ideophonic vocalizations, and embodied action such as gaze, facial expression, and gesture. By displaying emotions of different quality and intensity, and doing so with different configurations of semiotic practices, at different sequential moments, the participants show what they understand the current activity within the telling to be.

Keywords: emotion, stance, formulation, embodied action, interactional storytelling, conversation analysis

1. Introduction

“Have you ever been in a situation in which you were in serious danger of being killed?” In Sociolinguistic Patterns (1972), Labov reports that in answering this question, the respondent is “under some compulsion to show that there was a very real danger of his being killed; he stands in a very poor light if it appears that there was no actual danger. Often he becomes involved in the narration to the extent that he seems to be reliving the critical moment, and signs of emotional tension appear” (94). Labov goes on to describe the case of one respondent: “Within a few short sentences, a sudden and dramatic shift in his style took place. At this point, the speaker’s breathing became very heavy and irregular; his voice began to shake, and sweat appeared on his forehead. Small traces of nervous laughter appeared in his speech” (ibid.). And he observes that that the spontaneous storytellings of
another participant “show a remarkable command of pitch, volume and tempo for expressive purposes” (ibid.).

Labov inserted the “danger of death question” into sociolinguistic interviews as a methodological device to elicit spontaneous or casual speech. The need for such data was driven by a key goal of the variationist program, that is, to identify the phonological variables that distinguish between speech styles. The program had no interest in speech production features per se, let alone in nonvocal conduct. Yet Labov not only perceptively registered such features, as we have seen above. He also argued that “modulations of the voice production which affect speech as a whole” (ibid.:95) serve as evidence of speech style independent of phonological variables. Specifying these production features therefore provides a non-circular procedure to show how particular phonological variables are associated with a particular speech style. The method to specify such modulations, he recommends, “must follow the general procedure of linguistic analysis: the absolute values of tempo, pitch, volume, and breathing may be irrelevant, but contrasting values of these characteristics are cues to a differentiation of Style A and Style B. A change in tempo, a change in pitch range, a change in volume or rate of breathing, form socially significant signs of shift towards a more spontaneous or more casual style of speech” (ibid.:95, emphasis in original). To the four production features, he adds laughter as a fifth modulation, on the observation that laughter “is frequently heard in the description of the most dramatic and critical moments in the danger-of-death narration” (ibid.:95).

The danger-of-death question was asked by interviewers in the thoroughly interactional environment of the sociolinguistic interview, but it was not part of the variationist agenda to examine how the storytellings and the emotion displays that were so characteristic of them were organized as social interaction. In particular, beyond prompting the story by the question, the contributions of the story recipient to the telling were not examined. On the other hand, in the literature on the social organization of emotion in interaction, little attention has been given to fear – in contrast to studies of such emotions as surprise, disappointment, anger, anxiety, or distress (Peräkylä and Sorjonen 2012).

This chapter seeks to expand the small number of discourse-analytic studies on fear as an affective stance by examining how the participants in an ordinary conversation tell disaster stories in Japanese. Building on Labov’s speech production features in storytellings of life-threatening events and conversation-analytic work on emotion displays and emotion formulations, we ask how the participants recruit and combine different classes of multisemiotic resources, including language forms, cultural references, prosody, ideophonic vocalizations, and embodied action such as gaze, facial expression, and gesture, to bring off a disaster story as a frightening story. Drawing on Goodwin’s (2013) notion of lamination,
we want to see how both participants, teller and recipient, treat the story events as frightening, how the teller, as the main character in the story, produces herself and other characters as being frightened, and how the story recipient affiliates with the emotional stances that the telling practices convey.

2. **Resources and practices for the social organization of “being frightened”**

Presumably the emotional lexica of all languages include entries for “fear” or semantically related words and expressions, yet little is known about how such resources are used in emotion formulations in interaction. One exception is a study by Clancy (1999) on language socialization in mother-child interaction in Japanese. In three dyads of two-year-olds and their mothers, Clancy finds that the adjective *kowai* ‘be scary, be afraid (of)’ is the most frequently used lexical item to describe a specific emotion. The word is used to assess an action, object, or a person as frightening and to describe someone or something as experiencing being afraid. *Kowai* is therefore a prime resource for the socialization of fear as one aspect within the community’s affective culture. In mundane conversations with their children in domestic settings, the mothers model the use of *kowai* in specific interactional contexts and reinforce or reject the child’s use of it. Through repeated situated uses of the predicate in their everyday activities, the children learn to make culturally normative associations between the linguistic form, the emotion it formulates, and the persons, objects, and actions described as being or experiencing being *kowai*. Excerpt 1 is an illustration from a drawing activity.

**Excerpt 1. Obake (ghost) (adapted from Clancy 1999:1410)**

Yotchan is drawing a ghost and claims to have seen one.

1  Mother: aru? obake? mita koto aru?
    ‘You have? A ghost? You’ve seen one?’
2  Child: wao tte.
    ‘It goes “wao”.’
3 → Mother: “wao” tte yuu no? huun. kowasoo ne. yotchan kowakunai?
    ‘It goes “wao”? Gee. It sounds scary. You’re not afraid?’
4  Child: un.
    ‘No.’
5  Mother: huun. chugoi ((baby talk for sugoi ‘great’)) nee.
    ‘Wow. That’s great.’
6  Child: konosarenai.
    ‘I won’t get killed (by the ghost).’
7  Mother: korosarenai. (correcting pronunciation)
    ‘”I won’t get killed.”’
Responding to Yotchan’s “active noising” (O’Reilly 2005, see below) of the ghost, the mother describes the sound and by implication its imagined producer as scary, followed by asking the child whether he is afraid (3). Although Yotchan first denies and, upon the mother’s positive assessment, provides an account for why he is not afraid of the ghost, he then reverses his emotion avowal (8) and so claims the same feeling as the mother.

Fear formulations with kowai are also practices to control the child’s behavior, such as warning the child of danger and showing disapproval of unwanted conduct, as in Excerpt 2.

**Excerpt 2. Gun (adapted from Clancy 1999: 1410)**

Maachan (the child) is yelling loudly, pretending to fire a gun.

Child: ban!
   ‘Bang!’

Mother: a kowai na.
   ‘Oh, that’s scary/I’m afraid.’

Child: ban!
   ‘Bang!’

Here the mother claims to be frightened by the child’s unwanted action in an (unsuccessful) attempt to get him to stop the noisy play. By directing the child’s attention to the emotional impact that their conduct has on others, the mother encourages the child to see himself from other people’s perspectives. Clancy shows that the mothers use fear avowals and attributions, and respond to their children’s claims and ascriptions of fear with kowai, in a wide range of situations, including those involving third parties, fictional characters, animals, and inanimate objects. She also demonstrates that for the interactional organization of descriptions with kowai and other emotion predicates, the participants recruit generic sequence formats such as assertion – acceptance or assertion – rejection as well as question sequences. In this way the socialization of fear is brought off as part and parcel of ordinary everyday activities in the children’s and mothers’ domestic lives.

The excerpts cited from Clancy (1999) also illustrate the use of ideophones (wao, ban), a class of words that imitate sounds or give vocal shape to sensations. Ideophones have been defined as “words, not necessarily onomatopoeic, whose phonological structure itself encodes meanings” and that convey “dramatizations of actions or states” (Besnier 1990: 423). In a recent review, Dingemanse...
(2012:666) too highlights “the depictive mode of representation” as the defining
semiotic characteristic of ideophones. Functionally similar to represented talk
(Wilkinson, Beeke, and Maxim 2010), ideophones produce the represented scene
or object as an authentic vocal record while at the same time constructing the
speaker’s affective stance towards it. While kowai and other arbitrary emotion
words describe emotions, ideophones perform an event or state as emotional.
O’Reilly (2005) captures the functional similarity of represented talk and ideophones
by referring to their interactional use as “active noising” in analogy to
Wooffitt’s (1992) term “active voicing” for represented talk.

Research from linguistic anthropology suggests that ideophones are ubiquitous across languages (e.g., Dingemanse 2012), but they appear to be more prevalent in some languages than others. Commonly referred to as mimetics, ideophones are particularly widespread in Japanese discourse, although research has generally focused on their formal linguistic features (e.g., Akita 2010) rather than on what these vocalizations accomplish as discursive resources (but see Baba 2003; Kita 1997). Nevertheless the emotional dimension of mimetics is widely recognized, as Kita (1997) highlights by categorizing them as “affecto-imagistic.”

Only a small body of research has examined how ideophones are used in natural interaction, whether in Japanese or other languages. Exceptions include studies of ordinary conversation (Dingemanse 2012, 2013; Le Guen 2011), storytelling in a lecture (Szatrowski 2010), family therapy sessions (O’Reilly 2005), and talk with speakers who suffer from aphasia (Wilkinson et al. 2010). A converging finding of these studies is that ideophonic vocalizations regularly co-occur with gestural enactments (also Kita 1997). The contributions in Peräkylä and Sorjonen’s (2012) collection give much attention to the emotion-implicative production of linguistic resources through their temporal organization and prosody, yet no mention is made of ideophones (or onomatopoeia, “active noising”) as a vocal resource for constructing affective stance. Emotion displays with nonvocal resources, on the other hand, have been an expanding research topic. In addition to scrutinizing how emotional stance is produced through gesture and body movements (e.g., Goodwin 2007; Goodwin, Cekaite, and Goodwin 2012), recent studies put the spotlight on facial expressions that convey an emotional response to interactional events as well as objects and activities in the surround (Heath et al. 2012; Peräkylä and Ruusuvuori 2006, 2012). In storytellings, recipients regularly produce facial expressions that match the teller’s stance. Used on their own or together with brief response tokens, facial expressions allow recipients to affiliate with the teller’s posture while preserving the current distribution of discourse identities. Sugita (2012) examines specifically how participants indicate through nonvocal resources that they orient to an event as dangerous or threatening, and
how recipients simultaneously accomplish affiliation with the teller’s stance and structural alignment through facial practices.

In the study reported in this chapter, the participants utilize the types of resources described in this section for the telling of a disaster story in Japanese. We are particularly interested in examining how they locally orient to the story as a stance object by assembling language forms, other vocal resources, and nonvocal embodiments while coordinating their stance displays from their complementary discourse identities as story teller and recipient. For this project, Goodwin’s notions of lamination and transformation of substrates are helpful. Goodwin (2013: 12) shows how lamination (“a set of layers organized with reference to each other”) is a useful metaphor to see “a set of different semiotic fields organized as layers of diverse resources.” Actions and stances are built from arranging structurally diverse resources simultaneously and sequentially. A substrate is a configuration of resources at some interactional moment that is re-used later, either by another party or by the same actor. Transforming substrates involves the same fundamental operations as any form of repetition, that is, “(1) preserving structure provided by the activities of earlier actors while (2) systematically modifying that structure to build something new” (ibid.: 9). The transformation of substrates is fundamental in building a densely coherent form of publicly visible mutual orientation and as such critical to coordinating emotional stance in interaction.

3. Data

The data analyzed in this chapter come from a conversation between Peony and Yui (pseudonyms) at a café in Tokyo. Peony is an L1 speaker of Mandarin, a competent L2 speaker of English, and an intermediate L2 speaker of Japanese. At the time of this recording, she had lived in Japan for approximately a year. Yui is a colleague and friend of Peony and an L1 speaker of Japanese with some ability in English.

Over tea and dessert, Peony and Yui are talking about frightening experiences they have had while traveling and on other occasions. The analysis focuses on one of several stories in this thread.¹

¹. A note about romanization of Japanese. This paper generally follows the Hepburn style of romanization, with the exception of how long vowels are dealt with: phonemic long vowels are transcribed as double vowels (‘oo’, ‘ee’, etc.) unless the speaker pronounces them as a diphthong or glide.
4. Analysis

Emotion displays are not alike at different moments in a story telling. In fact, it is partly by displaying emotions of different kind and strength, and doing so with different configurations of semiotic practices, that participants show each other what they understand the current activity within the telling to be (Kupetz 2014; Selting 2010; Stivers 2008; Sugita 2012). The analysis is therefore organized according to the participants’ structuring of the telling into successive activities, “setting up the story as a disaster story,” “representing the disastrous events,” and “assessing the story as a disaster story.” The telling activities correspond to the general organization of story tellings into a preface sequence, a telling sequence, and a response sequence (Sacks 1974). The more specific descriptions highlight our analytical interest in seeing how the participants orient to the particular type of story as they move through the telling, and how they accomplish coordinated emotion displays throughout.

4.1 Setting up the story as a disaster story

One of the tasks that participants confront before the actual story telling gets underway is to work up a mutual sense of the kind of story that will be told. In addition to reference preparation, the story preface (Sacks 1974) is the dedicated place to also prepare the participants for the story as an object of emotional stance. In the conversation between Peony and Yui, a thread about frightening travel stories is already ongoing. In order to be recognized as a story in a series (Selting 2012), it is therefore the teller’s task to set up the upcoming story so that it fits into the ongoing thread and builds on the previous telling. The antecedent telling is a detailed story in which Peony describes how she had to sleep outside an airport in Hokkaido. Although both participants treat the story as frightening, Yui’s up-shot assessment focuses on Peony’s getting away unscathed and formulates her as *rakkii gaaru* ‘a lucky girl.’ That assessment prompts Peony to search for a story that ranks considerably higher on the fear scale.

Excerpt 3. [ARB-P130601] (see appendices for special conventions)

\[
\begin{align*}
p & \quad \text{+reaches for pen} \quad \text{+picks up pen & paper} \\
320 \quad P & \quad \text{+[demo(.) ima. (1.7) +ima} \quad \text{but now} \quad \text{now} \\
321 \quad P & \quad \text{+oboemashita.} \\
& \quad \text{remember-POL-PST} \\
& \quad \text{But, now...now I just remembered.}
\end{align*}
\]
But, New York, when I lived in New York.

You're kidding.
Peony launches the story as a “touched-off remembrance” (Frazier 2007) prompted by the active frame of frightening experiences (320–321). After locating the story in time and place (323–324), she writes “911” on a piece of scratch paper while saying the number in English2 (327). Although Yui is gazing at the paper,

2. For readers familiar with North American culture, “nine one one” (911) evokes the general emergency phone number in the US and Canada or the attacks on the World Trade Center on September 11, 2001. The standard English reference has become “nine eleven.” It is unclear why Peony, a longtime U.S. resident, does not use “nine eleven” for unambiguous reference, but it is unlikely that her reference selection is problematic for Yui. In Japanese, the most common reference to the September 11, 2001 events is “9∙11,” said kyō ten ichi ichi ‘nine dot one one,' or
the written and oral version of the number get the same neutral and minimal acknowledgment from her (329) as the time and place formulation (325), perhaps owing to there being no clear grammatical or intonational completion at this point. In the remaining 15 lines of the excerpt, the co-participants are occupied with achieving two purposes of the story preface, (1) to reach shared understanding that Peony was in New York during the events of September 11, 2001, and (2) to generate a shared stance towards the projected story as a disaster story. Orienting to the failed attempt to achieve mutual understanding, Peony repeats “911” while pointing her right index finger to her nose (331). This common self-referential gesture, re-used later in different versions (335, 339), connects Peony with “911” and places her at the scene. By repeating “911” rather than, for instance, substituting the metonymy with a description, Peony treats the name and its referent as shared knowledge and shows her expectation that Yui is able to recognize the referent. With the presumption of recognition, Peony orients to “911” as an event in recent world history and therefore shared public knowledge. To the gesturally elaborated repair, Yui responds by widening her eyes while simultaneously producing a confirmation check with rising intonation e-? >ita no<? ‘you were there?’ (332). With the lamination of facial expression, stance-marked response token, linguistically formulated candidate understanding, and prosody, Yui assembles a strong show of astonishment.3 The stance display presumes that Yui has recognized the reference “911,” consistent with Peony’s treatment of it as a shared epistemic and affective object. Peony’s emphatic confirmation through nodding and repeated confirmation tokens (333) is followed by an overlapping assessment in which Yui upgrades her affective stance by saying >uso.< ‘You’re kidding’ (334, 336), produced very quickly and in a staccato manner, while keeping her eyes widened. Peony reaffirms that she was indeed there by saying “yeah” in English and holding the paper on which “911” is written with her left hand while pointing to her shoulder with her right hand index finger (335). At the same time, Yui also opens and rounds her mouth (335 screen capture), a facial practice that, together with the widened eyes, constructs a recipient stance of astonishment (Heath et al. 2012; Selting 2010, see Note 3). Responding to Yui’s stance display, simply “911” kyu ichi ichi ‘nine one one’. Once Yui has adjusted her hearing to the codeswitch, she unproblematically recognizes Peony’s reference to ‘nine one one’. 3. Selting (2010) refers to similarly configured affect displays as “astonishment,” while Heath et al. (2012) call such exhibits “surprise.” In the ordinary English emotion lexicon, astonishment ranks higher than surprise and therefore appears to better capture the stance in question. However, we acknowledge that an analytically robust distinction would have to be grounded in systematic comparison of interactional data. The use of vernacular categories as tools for analysis is a considerable problem in this chapter and the wider conversation-analytic literature on emotion in interaction.
and in overlap with her repeated >uso.< (336), Peony begins to formulate her previous minimal confirmation with linguistic resources but abandons her turn when she runs into production difficulties (337–340). Yui’s and Peony’s next turns recycle the first response sequence (332–333) with some important transformations. After a gap of silence in which she may be waiting for Peony to finish her turn (341), Yui reuses her initial response *ita no ‘You were there?’ (332), but with an emphasized, elongated, and less highly raised format of ngːː. Here the construction can be heard as both completing Peony’s turn and seeking confirmation in a manner that conveys astonishment through its prosodic format in conjunction with Yui’s facial expression. Peony then confirms with the same layering of nodding and a series of vocal affirmation tokens, and sets down the paper that she has been holding. At the same time, Yui once more displays astonishment through the continued configuration of her face while producing a high-pitched vocalization with final rise (*hn↑nn?, 344). However in her next embodied action, Yui realigns her upper body from leaning forward toward Peony to a more relaxed posture. With these coordinated embodied actions, the participants mutually show themselves to be ready to move the story forward.

In the story preface, the progressivity of the talk is suspended after the teller’s initial announcement that she was in New York during “nine one one.” In terms of sequence organization, the story recipient uses multiple other-initiations of repair (“multiples”; Schegloff 2000) as a vehicle for displaying astonishment (Robinson 2006; Svennevig 2008) with vocal resources. During her own talk and while the teller is having a turn, the recipient exhibits her emotional stance through postural alignments and facial expression. The compounded effect of multisemiotic practices produces heightened emotional intensity towards the news delivery. By contrast, the teller sets the scene in a matter of fact way, ensuring the recipient’s understanding and supporting her emotional stance through reference preparation with multisemiotic practices but without conveying her stance towards the story or her part in it.4 Her project in the story preface is to get the recipient to orient to the story as highly tellable and newsworthy, and to take up an emotional stance that is congruent with the type of announced story. With sustained displays of astonishment, Yui treats Peony’s prefatory announcement in precisely the preferred way. The asymmetrical distribution of epistemic access is matched by complementary emotional stance displays.

4. The teller’s unemotional posture contrasts with Labov’s (1972) report on the intense emotional engagement seen in interview respondents’ danger of death stories. A contributing factor may be that in the sociolinguistic interviews, the stories are elicited by the interviewer, apparently without much forewarning, while in the story tellings analyzed here, the stories are self-selected by the teller and embedded in an ongoing series of frightening stories.
4.2 Representing the disastrous events

One particular feature of the telling sequence is that it has two culmination points, which we shall unimaginatively call “first climax” and “second climax.” It will become evident that the naming convention indexes not only the temporal succession of the apices but also the differential treatment they get from the participants.

As the telling moves towards the first climax, the participants’ coordinated orientations to the story as a disaster story shift significantly. The recipient engages many of the same multisemiotic practices as during the story launch, but in a manner tuned to the local evolvement of the telling. The teller, shifting away from the matter-of-fact, objectivizing mode of representation in the story preface, takes up a range of discursive measures that successively heighten the dramatic tension and construct the events as frightening.

In the course of setting the scene of the events further, Peony draws a picture of the Twin Towers that shows where her apartment was located in relation to them (Figure 1). The drawing becomes a significant semiotic resource in the following talk. We pick up the telling when Peony describes how she ran from the area together with her boyfriend and another friend.

![Figure 1. Peony’s drawing of the World Trade Center Towers](image)
Excerpt 4. [ARB-P130601]

3 people together run away-POL-PST

The three of us ran away together.

I see.

Yeah, we ran away.

When we ran away...TYHAAAH

Chapter 3. Disaster story 69
p  +GZ>paper
p  +RH pen>paper
y  +GZ>paper
519 P  +this+
   p  underlines drawing of buildings, then (GZ>Y
ey  GZ at paper (GZ>P
520  (1.3)
521 Y  u#so:[::#
   lie
   You’re kidding.
p  +nod
p  +LH up
522 P  [+nn nn nn.=
   Yeah yeah yeah.
523 Y  =nn::[:
   
   p  +BA mime ‘running’ +GZ>Y
524 P  +#°GHAAAA°# hh hh he he +.hh
   ONOMATOPOEIA
525  (0.4)
y  +GZ>paper
y  +RHIF>paper
526 Y  +e jaa [+saa
   well then
   So, then...
p  +GZ>paper
527 P  [hashi- [°rimash::ta.+°
   run- POL-PST
   We ran.

Excerpt 4 shows how the participants co-construct the first climax and the actions leading up to it as frightening events. Peony uses multisemiotic forms of representation to tell how the events developed temporally and display differentiated emotional stances toward them. She formulates her and her companions’ actions after they left their apartment with nigemashita (‘ran away’ or ‘escaped’), a verb associated with situations that are at the very least unpleasant, or (as in this case) frightening and dangerous. As she starts to say ni: (511), Peony simultaneously points to the drawing of the Twin Towers with her left index finger and turns her gaze to it, drawing Yui’s attention to it as a shared reference. In this way the
participants establish a joint attentional focus on the drawing and thereby locate
the characters’ running away in the scene. While they are gazing at the drawing,
Yui, pulling in her upper lip (511 screen captures), displays intense concentration
as Peony maintains a serious and alert facial expression throughout her turn. As
Peony says nigemashita, both she and Yui shift their gaze to each other while Pe-
ony keeps pointing to the drawing. With widened eyes and slightly opened and
rounded lips, Yui orients to the events as frightening. Yui’s visual actions and her
quietly produced acknowledgement (“soo nan da.”, 512) display affiliation with
the teller’s stance and alignment as a story recipient, yet Peony does not move the
story forward. Instead she reruns the description of the action with nigemashi-
ta. (513) while gazing intensely at Yui. Yui, for her part, keeps her gaze fixed at
the drawing while showing what Sugita (2012) calls “anticipatory fear,” a stance
fuelled by her knowledge of what the metonymy “911” entails.

Peony begins her next action (515) with a mispronunciation of nigeru, which
she self-repairs quickly. This builds into the initial turn format nigeru +toki ‘when
we ran away’ (517), which backgrounds the action described in the clause and
projects a climactic event. Peony completes the turn by representing a climax vo-
cally and visually (Hayashi 2003). Setting off the projected focal action from the
background, Peony pauses her speaking while raising both her lower arms and
hands. As she brings her arms and hands down, she simultaneously produces
the ideophonic vocalization TYHAAAH in what could be described as a “half
whisper,” that is, not particularly quiet but with little voicing and much aspiration.
With these laminated resources, Peony shifts her mode of representation from de-
scription to enactment (Wilkinson et al. 2010). Shifts as these are common prac-
tice in storytellings, where tellers use enactments such as represented talk to lend
authenticity and rhetorical impact to a climactic event (Holt 2000; Sidnell 2006).
Commenting on “active noising” as a form of enactment, O’Reilly (2005:749)
notes that such vocalizations “work to recreate the events as they happened at
the time, in the past, and allow visualization of reported events.” “Active noising”
represents the past event as something that the story recipient can witness in the
here and now. What becomes witnessable is not only the unfolding event but the
emotional attitude of the characters, while the teller simultaneously exhibits their
present stance towards the scene (Selting 2010).

These complex footings are evident in Peony’s multisemiotic representations
as well. The ideophonic vocalization and iconic gesture can be understood as con-
current and mutually elaborative enactments of the same event, possibly the fall
of one of the towers of the World Trade Center, which are depicted on the draw-
ing. During Peony’s enactment, Yui’s face takes on a terrified expression through
raised eyebrows, widened eyes and slightly opened lips while both participants
orient their head positions towards each other and hold mutual gaze. With their
coordinated actions, the participants accomplish a closely matching emotional stance (Couper-Kuhlen 2012). Pointing with a pen in her right hand and shifting gaze to the paper while saying “this” in English (519), Peony underlines the drawing of the buildings with several strokes. In this way she uses multiple resources of deictic reference to expressly link her immediately prior enactment to one of the buildings in the drawing. The multisemiotic practices achieve a dense representation of the referent as a stance object that gets affiliative uptake from Yui with *u#so:::# ‘you’re kidding’ (521). The lexical item *uso* is the same as in Yui’s displays of astonishment in Excerpt 3 but is here produced with a creaky voice and elongation of the second vowel. In response to the enacted catastrophe that has unfolded before Yui’s eyes, the assessment conveys a terrified stance.

Driving the story forward, Peony then uses an ideophonic *GHAAAAA* vocalization in the same mode of half-whisper as earlier (524), but simultaneously with a gesture that embodies a very different action: her arms pump up and down quickly while her face shifts downward, miming running with great force. Through the details of the enactment, the manner of the represented action gets encoded into Peony’s gesture (McNeill 1992), under Yui’s intense gaze. Shifting from performance to description, Peony’s turn in 527 *hashi-°rimash::ta.* ‘we ran’ can be understood as either formulating the enactment of the frantic running in 524 with linguistic resources or as completing the telling of the action, in which case line 524 is treated as an adverbial phrase constructed with nonlinguistic resources. With *hashiru* ‘to run’, Peony selects a verb that is not tied to dangerous situations in the same way as *nigeru* ‘to run away’, but the contextualizing gesture and vocalization provide a strong sense of running away from something dangerous.

After some unrelated talk, Peony reenters the story telling with an extended description of how she and her friends, not realizing what had happened, were “not very worried” (*shimpai shinai*) and “calmly” (*nonbiri*) walking and talking in the streets. Normalizing the scene in this way achieves what Sacks (1984: 419) called “the ‘nothing happened’ sense of really catastrophic events.” The description of the scene and the emotional tone of the telling, too long to document here, generate a strong contrast to the first climax as well as to the upcoming second. Starting with the contrast marker *demo* ‘but’ (not shown), Peony projects another turn of the events (Excerpt 5).
Excerpt 5. [ARB-P130601]

When we were walking, HYAacha (0.7) yeah.

Everyone quickly…

Peony builds the second climax in a manner similar to her production of the first (nigeru toki, (0.3) #TYAAAHAH# .hh ‘When we ran away (0.3) TYAAAHAH’, line 517), accompanied by the rising and dropping arm and hand gesture. In fact, it is through her reuse of the same structural resources that the performed event is immediately recognizable as a second climax (Goodwin 2013). But there are important transformations of the first scenario. One is that in the earlier instance, Peony sets the scene leading up to the climax as extraordinary and frightening, formulated in the backgrounding temporal clause When we ran/were running away. The climax then comes off as the culmination of an ongoing frightening scenario. In the current case, the backgrounding temporal clause arui:: (.) +toki ga a ‘When we were walking’ formulates the scene up to this point as normal, calm and not giving the characters reason to be fearful. The projected event now appears as highly contrastive. The contrast is taken further in line 590 when Peony starts the incomplete utterance minna hayaku ‘everyone quickly’ as she mimes running, albeit with less intensity than her previous mime (Excerpt 4). While the projected verb (nigeru ‘run away’ or hashiru ‘run’) is never produced and the utterance never completed, Peony’s recycled gestures again enact the posture of people who are frightened, as opposed to her earlier depiction of people walking and being unaware of what was happening. The method that Peony uses to work up the contrast resembles the format “I was just doing X … when Y” (Wooffitt 1992) or “At first I thought X, then I realized Y” (Jefferson 2004). As Wooffitt and Jefferson point out, the statement of a mundane “X” accomplishes a normalizing effect, or does “being
ordinary” (Sacks 1984). Against the backdrop of the X formulation, the Y formulation, such as accounts of President Kennedy’s assassination (Jefferson 2004) or accounts of the paranormal (Wooffitt 1992), will be understood as extraordinary. The normalizing account of the scene that frames the extraordinary event also describes the teller as someone who experiences the world as any normal person would and whose telling of the extraordinary is therefore trustworthy. Peony’s adaptation of the “I was just doing X ... when Y” format brings out the sense of acute fear in the enacted Y component as natural and credible.

Secondly, through several practices, the teller and recipient orient to the current episode as a second climax. Peony again shifts the mode of representation from description to enactment by laminating an ideophonic vocalization, produced in a similar manner as on the first occasion (i.e., with light voicing and aspiration) and an accompanying gesture. As a creative production HYAAAChaa can be taken to embody Peony’s subjective perceptual image of the sound. The gestural enactment also comes in a different version. The first time around, Peony depicted a falling object by lifting and dropping both of her lower arms and hands. The re-enacted gesture is a reduced version of the original, performed with the right arm and hand only. Lastly, vocalization and gesture are differentially coordinated. In the first instance, both were produced simultaneously, while in the second case, the vocal and gestural embodiments are produced consecutively, with the embodiment occurring during a 0.7-second pause after the ideophonic enactment. Although the second climax is still constructed as frightening, the different packaging downgrades the teller’s emotional stance towards it.

We also observe that Yui takes up a different recipient stance from the earlier occasion. During Peony’s delivery of the first climax, Yui’s gaze was fixed at Peony while her face showed a terrified expression through raised eyebrows, wide opened eyes, and an open mouth. Now Yui switches between gazing at the drawing and at Peony with a slightly open mouth that conveys intent recipiency but no fear. With her embodied actions, Yui’s recipient stance matches Peony’s less intense dramatization of the second climax. In this way, the participants collaborate in locating the disastrous events at different levels of emotional intensity. In the temporal structuring of the telling, the first event is told as the major climax, while the second event is told as the minor climax. The participants construct their understanding and emotional stance of the major and minor climax through methodic differences in the laminations of multisemiotic resources in the sequential contexts that they anticipate or respond to.
4.3 Assessing the story as a disaster story

While the telling of the story events was underway, the participants made extensive stance displays. In fact, some sequences we observed did little to advance the action in the story world. They were mainly dedicated to building the story as a stance object and to generate congruous recipient stances. Where the telling was concerned with developing the story line, the participants predominantly portrayed objects and events as frightening, and the characters in the storyworld and the conversational participants in the here-and-now as being frightened, through dense laminations of multisemiotic resources that generated the inference of fear. In contrast, explicit emotion formulations through lexical and grammatical resources were not used.

Excerpt 6. [ARB-P130601]

`y  GZ down
p  +pen taps on paper  +GZ>Y
664 P  +sono doki shin- shi+nu? (0.2) da to:
      that time die- die COP QT
665   (0.5)
666 P  [omoimashita.
      think-POL-PST
   At that time, I thought I was going to die.

y  +GZ>P
667 Y  [+omoimashita.
      think-POL-PST
   You thought.

p  +nod
668 P  +hai.
      yes
      Yes.

y  +GZ>paper
y  +slow nods
669 Y  +"nn:"[/::]

p  +GZ down  +GZ>Y
670 P  [shinu da to omoimashita.+ .hn  ko+waii: des.
      die COP QT think-PST-POL scary COP-POL

p  +GZ down
671 P  +"kowaii"
      scary
   I thought I was going to die. It’s scary. Scary.

672   (0.6)"
Alfred Rue Burch and Gabriele Kasper

7676

673 Y +tashika ni. =

definitely
definitely

674 P =+nn. +°TA TA TAAAAH

yeah ONOMATOPOEIA
Yeah. TA TA TAAAAH

675 Y +°taaa**+°

ONOMATOPOEIA
Taaa

676 (0.6)

677 Y +°sugoi ne.°

wow IP
Wow.

678 P °nn.°

679 (0.9)

680 P +min- (0.4) (gekkoo) +ee- +(0.4) +eega mitai;

everyone- (*rather) mov- movie seem
Everyone- it was like a movie.

681 (0.2)

682 Y +nn [eega mitai da [+ne/::¿

yeah movie seem COP IP
Yeah, like a movie, right?

683 P [n [da-
In the sequence following the telling of the story events, the participants collaborate in distilling what they make of the story and working up a joint assessment of it (Labov 1972; Sacks 1974). Their evaluation centers on the teller's emotional experience and the categorization of the events as terrifying. In this project explicit fear formulations figure prominently. At the outset, while tapping on the paper with the Twin Tower drawing, Peony leads the way by formulating her reaction to the events at the time they occurred, *sono doki shin- shinu?* (0.2) *da to: (0.5) omoimashita.*

5. ‘At that time, I thought I was going to die.’ With an overlapping collaborative co-completion as she is gazing at Peony (667), Yui shows empathetic understanding (Kupetz 2014). After mutual confirmation, Peony upgrades the upshot formulation by repeating the death avowal *shinu da to omoimashita*, followed by a characterization of the events as *ko+wai: des*°kowai° ‘It’s scary. Scary’. With the format of the *kowai* formulation, Peony accomplishes several things. First, by shifting from past to non-past tense she changes footing from speaking as a character in the story to speaking as teller at the present moment, which could adumbrate her readiness to move out of the storytelling. Secondly, the assessment is made as a categorical description that asserts the frightening character of the events as an objective fact and so views them as an object of socially shared stance (Clancy 1999) – in other words, anyone would experience events such as those described in the story as terrifying. At the same time, Peony displays her subjective stance through prosodic emphasis and vowel elongation while gazing at Yui and through a quietly said repetition while gazing downwards. In her upshot turn

---

5. The use of the copula *da* after a bare verb such as *shinu* ‘to die’ is prescriptively an error. However, there is no evidence here that it is oriented to as such.
in 670–671, Peony formulates and displays a coherent emotional stance as character in the storyworld and as teller in the ongoing conversation.

As response to Peony’s upshot formulation, an affiliative second assessment from Yui is due. However, following a gap of silence, Yui’s response tashika ni ‘definitely’ (673), while agreeing with Peony’s upshot formulation, does not display a congruent affiliative stance. After acknowledging the agreement, Peony pursues a more vigorous show of affiliation with a method that tellers commonly use to get a preferred response, that is by rerunning earlier elements from the telling (Mandelbaum 2013). Peony implements the practice by reenacting a version of the stronger first story climax from 517 with the same categories of semiotic resources, the gestural enactment and a non-conventionalized ideophonic vocalization, °TA TA TAAAAH°, produced in a half whisper. The re-enactment successfully generates an affiliative response: Yui reproduces the enactment in a reduced version, raising and dropping her right lower arm and hand while saying °taaaa° in an even quieter voice. After both participants direct their gaze at the Twin Tower drawing and so connect their co-enactments with the pictural representation of the scene, Yui shakes her head in a gesture of amazement while making the soft spoken verbal assessment °sugoi ne.° ‘wow’ that gets an agreement from Peony. With the adjacency pair, the participants treat the current phase of the evaluation as completed.

In her next verbal action, Peony initiates an upgraded assessment sequence that escalates in several steps. Using similes as a rhetorical device, she first compares the events to something that may be seen in a movie (eega mitai, line 680) and invites Yui’s agreement with a low final rise. Yui agrees by nodding and repeating the comparison with an upgraded version by ending her turn with the final particle ne/::¿. The particle indexes shared epistemic access and emotional stance (Cook 1992; Ishida 2009; Saigo 2011), the latter reinforced through the elongated production of the vowel and low final rise that corresponds to Peony’s turn-final intonation. As Couper-Kuhlen (2012) has shown, prosodic matching is a method for producing affiliation with vocal resources. By comparing a real-life event to something that is seen in movies, the participants align themselves as contemporaries in a shared popular culture. At the same time, the shared cultural knowledge is available to them as a resource for stance displays and for achieving a joint emotional perspective. Continuing to draw on the activated cultural store, Peony intensifies the assessment further with a reference to Godzilla (684). With the recognitional reference, Peony trades on Godzilla’s status as an icon of popular culture and therefore as an object that Yui will recognize. Slightly before saying the name, Peony starts moving her lower arms up and down in alternate rhythmic movements that can be seen to depict people running away from Godzilla as the creature wreaks havoc. Attracting Yui’s gaze, Peony further enacts people
running away with the ideophonic vocalization °KYHAAA°. With the naming and enactment of category-bound actions, Peony invokes Godzilla as an allegory for mass destruction. Overlapping with the vocalization, Yui claims recognition and affiliates with Peony’s escalated portrayal of the events as frightening with nodding and multiple repetitions of aa (685). Further into the sequence, responding to Peony’s depiction of people running in the streets, Yui reworks the movie and Godzilla imagery into an upshot formulation of her own, gojira ↑eega da↑ ne:: ‘It’s a Godzilla movie’ (not shown to conserve space). By making the assessment in first position and appropriating the lexical and rhetorical resources first used by Peony, Yui constructs a recipient stance that closely matches the teller’s.

As participants work up what they make of an emotionally intense story, they commonly shift from using practices that generate inferences to emotional stance to explicit emotion formulations with dedicated lexical and rhetorical resources (Prior, this volume; Rae 2008). This shift orients to the different interactional work that tellers and recipients have to get done during different phases of the story telling. While the telling is under way, a major challenge for the teller is to maintain the recipient’s interest in the story and evoke congruent affective stance(s). The recipient, for their part, is tasked with aligning to the telling as listener, that is, refrain from claiming full turns while displaying engaged listenership and stances responsive to those of the teller (Stivers 2008). As shown by Kupetz (2014) and Sugita (2012), such response practices can be effectively accomplished through visual devices only or through a combination of gesture and gaze with brief vocalizations. Collaborating on the take-away from the story, on the other hand, requires that such evanescent enactments are transformed into more robust formulations (Kupetz 2014). As we have seen, the participants draw on the Japanese emotion lexicon and rhetorical devices to formulate their stance on the story. At the same time, their interaction also shows that affective stance displays that are not of the preferred quality or strength prompt the teller to re-enter the story climax. In that endeavor the participants re-deploy the nonlinguistic resources from earlier in the telling.

5. Discussion

Previous literature has shown that types and strength of emotional stance, and the semiotic practices by which stance is achieved, are reflexively related to the ongoing activity. In story tellings, stance is one way for the participants to orient to the type of story. As story tellings evolve through successive components, stance is also co-constitutive of the narrative episode that is currently under way. During the telling of the 911 story, the participants oriented to the story as a
disaster story in distinctive episode-related ways. As the story telling was being set up, the participants’ project was to achieve mutual understanding that the type of upcoming story was the teller’s personal experience of a publicly known disaster. The newsworthy portion of the story launch was not the disaster itself but that the teller had personally experienced it. “911” served as a mutually accessible frame throughout the story, yet the participants oriented to their asymmetric epistemic status towards the teller’s personal experience of the event with complementary emotional stances. The teller (as knowing participant) announced the story in a factual manner, whereas the recipient (as unknowing party) oriented to the extraordinary character of the announcement with persisting astonishment. The recipient showed herself as being astonished with laminations of other-initiations of repair, facial expression, linguistic forms, and prosody. Each participant re-used their own but not the other party’s resources as substrates for later action and stance displays, such as Peony’s different versions of a self-referencing pointing gesture and Yui’s repetitions of uso ‘you’re kidding’ and her facial expression of astonishment.

In the telling sequence, the participants took up reciprocal stances both with regard to the quality and intensity of the emotion displays. The most noticeable change in telling practices was Peony’s shift from description to enactment when she constructed the dramatic culminations of the events. At these two moments, she assembled laminated telling devices that operated sequentially and simultaneously: (1) a temporal clause to back the pre-climatic scene, (2) ideophonic vocalizations, prosody, and iconic gestures to enact the fall of the Twin Towers, (3) iconic gestures and formulations to represent people running away. On both occasions, the configuration of resources represented the events as terrifying, but through several transformations of the original enactment (the substratum), the displays of danger and fear were weakened in the second episode. The recipient built reciprocal stances by reusing her earlier facial expressions and verbal actions as substrates that she upgraded as the first climax unfolded. During the second climax, by contrast, she downgraded her emotion displays through embodied actions that matched the teller’s stance. Finally, in the assessment sequence, the participants cooperated on formulating the take-away of the story. The teller shifted her mode of representation from resources that enabled inferences to her fearfulness to explicit formulations of fear with the lexical and grammatical resources of Japanese. When these assessments did not engender a sufficiently affiliative stance from the recipient, the teller pursued a reciprocal response by re-entering the telling of the climax ((2) above), using the earlier version as a substrate for yet another locally transformed re-enactment. In response, the recipient produced her own transformation of the embodied substrate with a return gesture (de Fornel 1992). Through their contiguous operations on the story climax as a substrate, the
participants achieved a congruent stance that allowed them to re-enter the assessment sequence. The teller further upgraded the assessment by using similes from popular culture, “like a movie” and “like Godzilla,” both of which were reused by the recipient in sequentially different but equally affiliative actions.

This study has shown that as tellers and recipients of a disaster story, an L1 speaker and an L2 speaker of Japanese achieve highly coordinated stance displays through multisemiotic resources and interactional practices that have been observed in story tellings and other social activities between first language speakers. The study confirms that for the construction of emotion in talk, participants draw on generic categories of resources and practices that they configure in context- and activity specific ways to produce specific kinds of emotion and different degrees of emotional intensity. This general upshot from the analysis raises the question whether the stance-related practices are in any way specific to second language talk.

Compared to proficient speakers’ talk, some grammatical and pragmatic resources are absent in Peony’s use of Japanese. One such class of features are interactional final particles such as ne or yo, which are common resources for displaying stance and often attached after a predicate. Interactional particles afford the story recipient with a cue how to respond in an appropriately affiliative manner. These particles have been found to be challenging for learners of Japanese (Ishida 2009; Saigo 2011), and their absence in environments where they are normatively expected can be consequential for the interaction (Saigo 2011). In other interactions in the data corpus, Peony’s co-participants do sometimes orient to such absences. Yet as recipient of the 911 story, Yui does not show – through delays, other-initiations of repair, or “misunderstandings” – that her understanding suffers.

Another component that is sometimes not included in Peony’s turns is a grammatical predicate. For example, in Peony’s story preface in Excerpt 3, her turn with the utterance “>nine one one<” (line 327), spoken without noticeable falling or rising intonation, is grammatically, prosodically, and pragmatically incomplete (Ford 2004). Similarly, Peony’s turn that enacts the crashing of one of the Twin Towers Excerpt 4 (line 517) requires a predicate to be complete. In both of these instances, the gaps of silence after Peony’s turn, and Yui’s neutral minimal acknowledgement token in Excerpt 3 (line 329), suggest that the turns provide Yui with insufficient cues how to respond. Here the normative grammatical predicate is noticeably absent for the co-participant, and its absence is interactionally consequential.

However in the disaster story telling, these two are the only instances in which Peony’s non-use of a grammatical predicate becomes visibly problematic for the story recipient. All other absences and unusual features in Peony’s talk may be registered by an overhearing observer, but they do not have consequences for
Alfred Rue Burch and Gabriele Kasper

the organization of the talk. Rather Peony artfully engages her recipient by dramatizing the story with complex multisemiotic configurations of stance marking practices. With these methods Peony accomplishes emotional stance in ways that render the absence of standard grammatical and pragmatic devices interactionally inconsequential.

References


Prior, Matthew. T. (this volume). “Formulating and Scaling Emotionality in L2 Qualitative Research Interviews.”


**Appendix A: Glossing conventions**

<table>
<thead>
<tr>
<th>TOP</th>
<th>Topic marker</th>
<th>PST</th>
<th>Past tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUB</td>
<td>Subject marker</td>
<td>POL</td>
<td>Polite suffix</td>
</tr>
<tr>
<td>LK</td>
<td>Linking particle</td>
<td>NOM</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>OBJ</td>
<td>Object marker</td>
<td>IP</td>
<td>Interactional particle</td>
</tr>
<tr>
<td>QT</td>
<td>Quotative marker</td>
<td>*</td>
<td>Unglossed particle</td>
</tr>
<tr>
<td>Q</td>
<td>Question</td>
<td>*word</td>
<td>Non-targetlike form</td>
</tr>
<tr>
<td>COP</td>
<td>Copula</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix B: Special conventions for representing embodiment

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>hand(s)</td>
</tr>
<tr>
<td>$F$</td>
<td>finger</td>
</tr>
<tr>
<td>$R$</td>
<td>right</td>
</tr>
<tr>
<td>$IF$</td>
<td>index finger</td>
</tr>
<tr>
<td>$L$</td>
<td>left</td>
</tr>
<tr>
<td>$2Fs$</td>
<td>index and middle fingers</td>
</tr>
<tr>
<td>$B$</td>
<td>both</td>
</tr>
<tr>
<td>$3Fs$</td>
<td>index, middle, and ring fingers</td>
</tr>
<tr>
<td>$GZ$</td>
<td>gaze</td>
</tr>
<tr>
<td>$+$</td>
<td>place where action begins, description of action</td>
</tr>
<tr>
<td>$+$</td>
<td>place where action begins in relation to talk</td>
</tr>
<tr>
<td>$/$</td>
<td>stroke or beat of gesture</td>
</tr>
<tr>
<td>$---$</td>
<td>holding gesture or gaze in place</td>
</tr>
</tbody>
</table>

**Palm positions**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Position</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRD</td>
<td>prone down</td>
<td>(palm facing ground)</td>
</tr>
<tr>
<td>SPU</td>
<td>supine up</td>
<td>(palm facing up)</td>
</tr>
<tr>
<td>SPV</td>
<td>supine vertical</td>
<td>(palm facing in)</td>
</tr>
</tbody>
</table>