

Salientizing the breaks in talk: A study of Japanese segmentizing

Emi Morita

Department of Japanese Studies, National University of Singapore, Singapore
jpsme@nus.edu.sg

Abstract

In naturally occurring conversation, Japanese speakers often break up their turns at talk with seemingly random or disfluent pauses that break the flow of talk into a series of successive small segments which may not be semantically coherent. Moreover, the boundaries between such segments are often made salient via the attachment of interactional particles, such as *ne* and *sa*. Empirical observation of such naturally occurring partitioning of talk reveals that such “semantically irregular” segmentation is used by both speakers and their recipients to accomplish a legitimate communicative function in managing the fine-tuned choreography of moment-by-moment conversational interaction.

Index Terms: utterance segmentation, interactional particles, Japanese conversation

1. Introduction

In naturally occurring conversation, speakers often break up their turns at talk with seemingly random or disfluent pauses, glottal stops, and other perturbations that break the flow of talk into a series of successive small segments. In this regard, the production of conversational talk by Japanese speakers appears to be considerably more segmented and “fragmentary” when compared with the talk of English speakers. Maynard [1], for example, has reported that the average length of what she terms “pause-bounded phrasal units (PPU)” is 2.36 words in Japanese – compared to about 4.00 words per equivalent unit in English. In Japanese conversation, such segmentations are often quite prosodically salient as they are often marked with, what I call interactional particles such as *ne* and *sa*. Figure 1 shows the pitch track of an utterance that indicates how such segmentation is hearably produced by a speaker. Here each subunit has been made saliently distinct not only by the use of a significant pitch raise, but also by the simultaneous marking with an interactional particle (in this case, the particle *ne*).

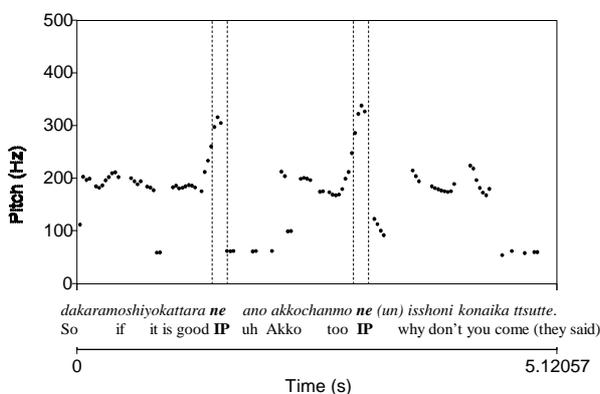


Figure 1: Pitch track of an utterance with segmentation marked with interactional particle *ne*

In their prosodic aspect, such hearably distinct “talk segments” have been referred to as Intonation Units (IU), defined as a “stretch of speech occurring under a single unified

intonation contour” [2]. According to Chafe [3], each single IU limits itself to expressing no more than one new idea. Thus defined, an IU is considered to be a reflection of the cognitive processing limitations on the part of speakers and their listeners. A more interactional view of this phenomenon, however, proposes that these breaks specifically draw attention to the segmental nature of each unit of information in the course of the linear production of a complete utterance, so that the recipient may act upon each individual piece of information presented in real-time. For example, Fox, Hayashi and Jaspersen [4] propose that Japanese speakers produce relatively short PPU in order that recipients may acknowledge or question the elements of a speaker’s turn-in-progress without waiting to discover how the final product of that turn will be constructed. Similarly, Nagahara and Iwasaki find functional commonalities between pitch excursions at the end of an utterance and at utterance-medial positions, since both “solicit the involvement of the interlocutor” [5].

However, naturally occurring Japanese conversational data reveals that this way of speech production may not merely be a reflection of the bit-by-bit manner of “information” transfer. For not only are such segmented parts often semantically “empty” as a particle attaches to a so-called “hesitation marker” or “filler” such as “*ano...*” (the equivalent of English “uh...” or “well...”), but such interactional particles may be inserted in the middle of a noun phrase as well (as we shall see in the data which follows), deliberately breaking up an otherwise tightly connected semantic unit.

The questions that I set out to ask myself in this study are as follows: Why would a speaker deliberately break up an utterance in a way that makes disjoint an otherwise coherent semantic unit? What factors might contribute to the decision of a conversationalist to “break up” such talk at such unlikely places? How do the speaker’s listeners react to such segmentation? If it is semantic constituency that is being broken up by such segmentation, rather than distinct packages of “information”, what is it that speakers who are doing such segmentizing are demarcating?

This study argues that such segmentation, and its salientizing not only by pitch contour but also with the addition of interactional particles, functions as a conversational resource that allows both speakers and listeners to deal with immediate conversational problems or concerns as may arise in the course of an utterance production. The study concludes that such segmentation is indeed not merely attributable to speech “disfluency” or idiosyncratic “quirkiness” - but is, instead, an integral component for effective communicative practice, and a resource for conversationalists to fine-tune the moment-by-moment contingencies of real-time conversational interaction.

2. Method

2.1. Conversation Analysis

In the attempt to understand the production of talk as collaborative activity between speakers and their listeners, I use the methodology and analytical approach of Conversation

Analysis (CA) to explicate the “interactional practices” in which the conversationalists under study show themselves to be engaged.

Studies of CA reveal that linked actions in adjacent positions are the sequential basis for the construction of intersubjectivity as a structural resource [6, 8]. It further holds that the ways in which conversation develops and that a given stretch of talk is understood is objectively analyzable not only for the participants of the talk, but also for conversational analysts with access to recordings of the data. Moreover, CA focuses on not only the verbal aspect of conversational talk itself, but also the physical environment where the conversationalists are situated, allowing us to see multi-modal nature of conversational interaction [6], which will be an important aspect of this study, as I attempt to qualitatively analyze the phenomenon of speakers’ segmentation of their talk in naturally occurring conversation.

2.2. Data

My data consists of audio-video recordings of naturally occurring conversation between native speakers of Japanese. In this paper, I will present just two short excerpts from my corpus [7]. One is a mother-child interaction and the other is a conversation between two female friends. The keys to the transcription notation and Japanese gloss symbols used in my analysis appear in section 6.

3. Segmentizing talk

3.1. Findings from Studies in Conversation Analysis

Studies in Conversation Analysis have long noticed recurrent instances of utterance-internal “breaks” and have revealed many of the discourse functions associated with them. Jefferson [8], for example, has noted that “perturbation” at various unit-initial positions may serve to index distinct activity junctures; Goodwin [6] has powerfully demonstrated how “phrasal breaks” may serve to attract the recipients’ eye-gaze, and Sacks & Schegloff [9] have demonstrated that English speakers may mark the proper name of a person who is mentioned for the first time in the conversation with an upward intonation, in order to solicit a recipient’s signal of recognition without disturbing the whole turn’s action.

Creating utterance-internal segmentation appears to be a methodical conversational resource in Japanese as well, as I will discuss in more detail in the following section.

3.2. Japanese interactional particles as a linguistic resource for salientizing segmentation

Japanese speakers break up their turns at talk into smaller hearable sub-units by inserting prosodically salient breaks within a single turn at talk. Notably, such breaks are often accompanied by interactional particles such as *ne* and *sa*.

These interactional particles allow Japanese speakers to demarcate segments even at the sub-phrasal level. With the ability to insert interactional particles at any point in the production of talk, Japanese speakers are able to break up – and thus effectively segmentize – even a single semantic constituent, as illustrated below:

Excerpt (1) [M&H]

H: (.4) *e, sono sa,*
INTJ that PRT

Kankoku no otoofu tte iu hanashi wa=
Korean ATT tofu QUO say story TOP

=uwasa o miminishita n da kedo
rumor OBJ hear:PAST SE PRED but

‘e, that *sa,* rumor about the Korean tofu, I heard about but...’

In the above example, the adnoun *sono* plus the Noun Phrase *sono kankoku no tofu tte iu hanashi* (‘that rumor about the Korean tofu’) – which together form a semantically and syntactically tightly connected constituent – is divided into two units by the interactional particle *sa*. The following diagram makes this point clear:

<i>Sono sa</i>	<i>kankoku no tofu tte iu hanashi wa</i>
That IP	Korean ATT tofu QUO say store TOP
‘that	rumor about the Korean tofu’

Breaking a phrase in such a phrase-medial position would indeed be odd, were the function of segmentizing only to mark some stretch of *propositional* information to which the recipient should pay attention. Yet, my data shows that such inter-phrasal “breakings” are not uncommon in spoken Japanese.

In an effort to explain this phenomenon, I will now examine two similar instances of naturalistic conversational data in order to demonstrate how such particle-marked segments explicitly indicate the collaboration-necessary nature of the marked, ongoing conversational “move”. Close analyses of the sequence-by-sequence development of talk will reveal how such “marked breaks” work to highlight the talk-so-far as an interactionally relevant move, and how conversationalists are oriented to such minutely but explicitly marked breaks.

4. Analysis

As with most of what goes on in conversation, the segmentizing of talk is not merely a reflection of a speaker’s independent mental planning of their turn’s design. Rather, such behavior may also be shaped by elements in the immediate interactional environment, such as the physical distance to the object of conversation, or to another person’s position in space. The next example shows how participants’ physical positioning can affect the segmentation of talk.

Excerpt (2) [Shohei]

(Child and Mother are at the table, and mother is making origami.)

1→C: [*ne, kore sa:*
IP this IP
‘look, this *sa-*,’
(((starts moving his upper body and arm toward the origami))

2 M: *un.*
‘uh-huh’

3 C: [*kasa ni mo chotto nitenai?*
umbrella to also little resemble-NEG
‘doesn’t (this) look like an umbrella too?’
(((grabbing the origami))

4 M: *kasa ni niteru ne[:.*
umbrella to resemble IP
‘it does look like an umbrella, doesn’t it?’

5 C: [*un.*
‘Yeah.’

In this segment, the child is offering a candidate assessment using the grammatical form for seeking an agreement. Within this action, he inserts the interactional particle *sa* after a

demonstrative pronoun *'kore'* (this) demarcating it from the rest of the action. Why does he demarcate his talk here?

When we look at the physical environment, as Figure 2 shows, the child has not yet reached the origami at the onset of his turn (Line1). The mother is also holding another origami paper; hence, she is looking down. Thus, although he is saying *'kore,'* the recipient of the turn (mother) cannot see what he is actually pointing to (Figure 2). But by attaching *sa* to a demonstrative pronoun, he is able to effectively highlight that he is “pointing” at something right now, and in so doing, he creates an interactional space for the mother to notice this move. Indeed, the mother does respond to this segment with *'un'*. This short response was minimal, but suffices to indicate that the mother has acknowledged the child’s move of pointing and is now showing her readiness to receive the rest of the action. The response to such a recognized “segment” is also consequential, in that the mother is now committed to finding out what he is pointing to. Only after receiving this minimum acknowledgement by which the child’s move is *acknowledged*, does the child continue the rest of his talk as his hand is physically reaching the object.

During the course of the child’s production of the utterance “looks like an umbrella too, doesn’t it?”, the mother also moves her eye gaze to the origami that he is holding (Figure 3).



Figure 2: “ne, this sa”



Figure 3: “looks like an umbrella too, doesn’t it?”

By the time his action of commenting on the resemblance of the Christmas tree to an umbrella is completed, the mother is looking at the object. For the child to successfully accomplish this action, gaining his mother’s eye gaze to the object was a prerequisite. Such pointing of an object must be accomplished interactively, and by inserting an interactional particle and segmentizing the linguistic element of the demonstrative pronoun, the child successfully accomplishes his “pointing” even before he completes his utterance.

What the child is primarily segmenting here is his *action* – as opposed to the propositional content of his utterance – and the manner in which he does so demonstrates how segmentation is carefully calibrated in the phenomenal environment of participants within the temporal unfolding of each other’s position while the action is proceeding. Specifically: he has successfully dealt with a contingency problem of starting a different action – along with problem in his physical environment (his mother’s not looking appropriately) - by shaping his utterance in a cumulative and progressive manner and by highlighting the ongoing “move” of *pointing* as he accomplishes a larger level of action of “offering a candidate assessment.”

My second examples shows how a speaker segmentizes her talk to deal with a potentially problematic conversational move of “changing topic” in the course of formulating a question. In the next excerpt, graduate students Midori and Hiroko are talking about what is good to eat in Los Angeles. As a senior colleague who lived in Los Angeles for long time, Midori is listing some of her favorite things to eat there. As the excerpt below starts, Midori verbally indicates that she is still thinking of more options.

Excerpt (3) [M&H]

- 1 M: *ato wa doko: ga e< ii ka na:*
Other TOP where NOM FRG good Q IP
‘Where else would be good, I wonder...’
- 2 (4)
- 3→H: *e, sono sa,*
INTJ that IP
- 4 *kankoku no otoofu tte iu hanashi wa*
Korean GEN tofu QUO say story TOP

=uwasa o miminishita n da kedo
rumor ACC heard SE PRED but
‘Eh, that *sa* rumor about the Korean tofu, I heard but...’
- 5 M: *un!*
‘Yeah.’
- 6 H: *sore tte nani?*
that QUO what
‘What is it?’
- 7 *o-. ni- nihon no otoofu to °onaji na no?*
FRG FRG Japan GEN tofu as same PRED:ATT SE
‘Is it the same as Japanese tofu?’

In line 3, where Hiroko starts her new turn, she interpolates an interactional particle, *sa*, after the adnoun *'sono'* (that) breaking up a semantically coherent noun phrase in the fashion discussed in section 3. If segmentation were for “conveying information in a bit-by-bit fashion”, this would be an odd place to do so. The discourse deictic term *sono* here is an adnoun that has the sole grammatical function of modifying a noun. In other words, *'sono'* alone cannot appear as an independent word - it must always appear with a noun. In the above extract, only when *sono* is together with its modified part *'kankoku no tofu tte iu hanashi'* (‘rumor about the Korean tofu’), does the segment constitute a semantically coherent unit. Why is the speaker highlighting the adnoun *'sono'* alone here? In order to understand this, we need to pay attention to the sequential environment.

When Hiroko is about to bring the conversation back the previously mentioned item (“Korean tofu”) in the current

discourse, there are several outstanding issues that must be dealt with. While Midori displays out loud what she is thinking in line 1, her eye gaze is not directed at Hiroko - rather, her gaze is directed to the middle distance. In this way, Midori displays that she is not selecting anyone as the next speaker. Midori is indicating that she is now engaged in searching for a new food item to talk about – so in a way, she is reserving the next speakership and also indicating that the next food item will be proposed by her. Midori is thus systematically not selecting Hiroko as the next speaker and is still holding the floor even after this turn, during the silence.

So now, when Hiroko wants to bring back the conversation to the previously mentioned item, it has to be done in such a way that she is doing so “acknowledgingly” of what Midori is doing at this moment. By salientizing the adnoun *sono*, Hiroko demarcates it from the rest of the talk with the interactional particle *sa* (for other specific indications marked by *sa*, see Morita [7]), and by so doing, Hiroko indicates the problematic nature of the move up front (i.e., at the earliest point of her utterance). Rather than just advancing this disjunctive move as one seamless action, by segmentizing *sono*, Hiroko is letting Midori know that this move may be “backward looking” (as anaphoric term *sono* points to some referent previously appeared in the discourse), and therefore “disjunctive” in relation to on-going talk. This, in turn, works as a warning for Hiroko’s problematic action on-the-way, and also works to display of her awareness of the potentially problematic nature of her own move. Hiroko’s use of segmentation at this precise point is indeed meaningful and contextually motivated.

Such a design of the turn, i.e., highlighting a move within an action, can be seen to be meaningful for sociopragmatic reasons. At the sequential conjunction where the other participant is engaged in a totally different activity, to highlight one’s possibly-problematic move as soon as possible indicates the speaker’s awareness of the contingency problem that might thereby come in play.

Significantly, the Midori’s subsequent response shows that Hiroko’s move here is both understood and accepted as such, as indicated by Midori’s continuer in line 5. This example shows that while such segmentation may not constitute a semantically coherent constituent *per se*, Japanese conversational grammar does allow speakers to break up such phrases in order to highlight a particular move *within* a larger coherent conversational action.

5. Summary

The Japanese conversational data that I have presented here suggests that the deliberate segmentizing of an utterance by a speaker may not be so much a product of “disfluency” or of a making of the “information value” of the proposition contained within its boundaries, but more likely is a resource for making intelligible what exactly is happening in the course of an ongoing conversational action. By highlighting that particular “part of the talk” so marked, participants can monitor and negotiate each others’ in-progress conversational moves. Such segmentations are thus carefully designed to include the necessary lexical items as being recognizable to do this specific job and to signal that this sub-unit formulation is not random. (Here, the those lexical items are *kore* and *sono*, though the phenomena is in no way limited to deictic terms, as is evidenced in data that I have not discussed here.) Such examples show that “referencing” can be problematic at certain junctures of the conversation, and that when recognized as such, speakers can create a negotiation space by demarcating such “moves” with interactional particles.

Moreover, my data [7] indicates that such salientized segmentation often appears at a quite early point in the turn construction. As the potential contingency problems are something that needs to be indicated and dealt with as soon as possible, the occurrence of such segmentizing “breaks” toward the beginning of an utterance makes sense.

In short, a CA analysis of the above two examples suggests that such “semantically irregular” segmentation is used by both speakers and their recipients to accomplish a legitimate communicative function in managing the fine-tuned choreography of moment-by-moment conversational interaction.

6. Gloss symbols and transcription convention

Japanese gloss symbols abbreviations:

ACC = accusative	NOM = nominative
ATT = attributive form	PRED = predicate formative
FRG = fragment	Q = question
GEN = genitive	QUO = quotative
INTJ = interjection	SE = sentence extender
IP = interactional particles	TOP = topic marking
NEG = negative	

Transcription notation:

[]	overlapping talk
(0.5)	length of silence in tenths of a second
(.)	micro-pause
.	falling intonation
?	rising intonation
!	an animate tone
,	continuing intonation
=	contiguous utterances (no break or gap)
:	prolongation of immediately prior sound
°°	talk is quieter than the surrounding talk
(())	transcriber’s descriptions of events
→	points out the phenomenon under examination

7. References

- [1] Maynard, S. K. Japanese conversation: self-contextualization through structure and interactional management. Norwood, N.J.: Ablex. 1989.
- [2] Du Bois, J. W., Schuetze-Coburn, S, Cumming, S. & Danae, P. Outlie of discourse transcription. In J. Edwards and M. D. Lampert (eds.), Talking data: transcription and coding in discourse research, 45-89. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers. 1992.
- [3] Chafe, W. L. *Discourse, consciousness, and time: the flow and displacement of conscious experience in speaking and writing*. Chicago and London: The University of Chicago Press. 1994.
- [4] Fox, B. A., Hayashi, M, and Jaspersen, R. Resources and repair: a cross-linguistic study of syntax and repair. In *Interaction and grammar*, E. Ochs, E. A. Schegloff and S. A. Thompson [eds.], 185-237. Cambridge: Cambridge University Press, 1996.
- [5] Nagahara, H. & Iwasaki, S. 1995 “Tail pitch movement and the intermediate phrase in Japanese”. Unpublished manuscript
- [6] Goodwin, C. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press. 1981.
- [7] Morita, E. *Negotiation of contingent talk: The Japanese interactional particles ne and sa*. Amsterdam: John Benjamins, 2005.
- [8] Jefferson, G. “Sequential aspects of storytelling in conversation”. In J. Schenkein (Ed.) *Studies in the organization of conversational interaction*. New York: Academic Press, 219-248, 1978.
- [9] Sacks, H, & Schegloff, E A. Two preferences in the organization of reference to persons to their interaction. In G. Psathas [Ed.], *Everyday language: Studies in ethnomethodology*, 15-21. New York: Irvington. 1979.