

Anti-zero pronominalization: when Japanese speakers overtly express omissible topic phrases

Yasuharu Den¹ & Natsuko Nakagawa²

¹ Faculty of Letters, Chiba University, Japan

² Graduate School of Human and Environmental Studies, Kyoto University, Japan

Abstract

In this paper, we focus on cases where Japanese speakers overtly express a topic phrase that could have been omitted. We call this phenomenon *anti-zero-pronominalization* and hypothesize that this helps speakers gain time for planning a following utterance; anti-zero-pronominalization is another option to deal with cognitive load at the beginning of an utterance in addition to fillers and other speech disfluencies. Based on a quantitative analysis of a corpus of spontaneous Japanese dialogs, we investigate the difference between overt topic NPs and zero-pronouns. We show that i) the utterance is more complex when the topic is expressed as an overt NP than when it is expressed as a zero-pronoun; ii) turn-initial items such as fillers are produced less frequently when overt NPs appear than when zero-pronouns appear; and iii) the utterance becomes more complex when the last mora of the topic is more prolonged.

Index Terms: zero-pronouns, topic phrases, cognitive load, Japanese dialogs

1. Introduction

In Japanese, speakers can omit arguments of predicates when they are recoverable from the context. Unlike English, in which ellipsis is possible only in limited syntactic configurations such as coordinated structures, Japanese has rather weak constraints on the use of ellipsis, or *zero-pronouns*. Not only subjects but also objects and other arguments can be omitted regardless of syntactic configurations. Semantic constraints are also weak; in addition to first and second person pronouns, expressions referring to third persons and non-animate or abstract objects are subject to zero-pronominalization.

There have been a number of studies on Japanese zero-pronominalization in terms of pragmatic constraints of the use of zero-pronouns, models of zero-anaphora resolution, characteristics of zero-pronouns in real data, and so on [1–6]. Based on quantitative analyses of narratives and conversations, some researchers reported that in spoken Japanese about 70% of arguments in their data were zero-pronouns [2, 3, 6]. According to some scholars [7, 8], the choice between overt NPs and zero-pronouns is crosslinguistically determined by the cognitive status of the referent in question. In general, zero-pronouns are used when the referent is assumed to be activated in the hearer’s mind, while overt NPs are used when the referent is not assumed to be activated.

There are, however, some cases in which speakers overtly express arguments of predicates even though the referent can be assumed to be activated. For instance, consider the following exchange:

(1) L: zyosee-wa i-nai-n-desu-ne
 women-TOP be-NEG-N-POL-FP
There were no women, right?

R: zyosee-wa-ne i-masi-ta-kedomo (0.4)
 women-TOP-FP be-POL-PAST-though

mata ryoo-ga tigat-te
 as well dormitory-NOM be.different-and
There were women, but the dormitories were different.

(D01M0047:403.854-408.814)

In responding to L’s question, R could have omitted “zyosee (*women*)” and simply said, “i-masi-ta-kedomo, ... ((Yes,) *there were, ...*),” which is perfectly acceptable in this context because the expression “zyosee (*women*)” is mentioned in the immediate context. The speaker, however, repeats the topic phrase “zyosee-wa-ne.” We call this phenomenon *anti-zero-pronominalization*.

One possible motivation for speakers to anti-zero-pronominalize omissible topic phrases is to gain time for planning a following utterance. By overtly expressing a topic phrase at the beginning of an utterance, the speaker can delay the production of the substantial content of the utterance. In this respect, we hypothesize that anti-zero-pronominalization has a similar function to speech disfluencies such as fillers [9, 10], word repetitions [11, 12], and prolongation [13].

In this paper, we elucidate factors behind the use of anti-zero-pronominalization in Japanese dialogs. In particular, we test hypothesis that overt topic phrases, as opposed to zero-pronouns, are used when the speaker’s cognitive load in speech planning is relatively high. To test this hypothesis, we first extract from a corpus of spontaneous dialog exchanges like (1), where a discourse entity overtly expressed in the first utterance is repeated in the second utterance by the other speaker and expressed either as an overt NP or as a zero-pronoun. We then examine whether or not the expression type (overt vs. zero) of the topic affects some production-related variables such as the duration of the substantial part of the utterance and the presence of prefaced items at the beginning of the utterance.

2. Methods

2.1. Corpus

We used the dialog subset of the core data of the *Corpus of Spontaneous Japanese* [14]. The data consists of 18 sessions of dialogs produced by 6 dyads, each of which participated in 3 different kinds of sessions: SPS-Int, APS-Int, and Task. In the SPS-Int dialog, one of the participants of each dyad, called ‘interviewer,’ interviewed the other participant (‘interviewee’) on the simulated public speech (SPS) the interviewee had given before the interview.

L: sono sanzyoo-no heya-ni hokani-wa donna kadenseehin-toka oi-te-ta-n-desu-ka
 that three.tatami-GEN room-in other-TOP what.kind home.appliances-like have-PROG-PAST-N-POL-Q
What kind of home appliances did (you) have in that three-tatami room?

R: | a | (0.2) | sore-wa | kekkoo odoroku-hodo takusan hait-te |
 | oh | | that-TOP | very surprising-as alot had-and |
 | **Preface** | | **Topic** | **Body** |
Oh, (I) had surprisingly many of those things.

(D01M0019:225.268-231.874)

Figure 1: The structure of the responding utterance.

In the APS-Int dialog, the interviewer and the interviewee talked about the academic presentation speech (APS) the interviewee had given before the interview. In the Task dialog, each dyad performed a task-oriented dialog without roles of interviewer and interviewee.

The core data of the corpus is annotated with rich information, including phonetic segments, words, phrases, and clause units, and provided in the form of relational database [15]. Starting and ending times of units at each level of granularity are precisely identified, which enables us to conduct detailed analyses based on time information.

2.2. Annotation

We first extracted exchanges between two interlocutors, where a discourse entity overtly expressed in the first utterance was mentioned in the second utterance by the other speaker in the form of either an overt NP or a zero-pronoun. The detailed procedures are the following:

1. We identified clause units, which we regarded as utterances [16]. We combined two clause units that were temporally proximate if the first unit was response tokens, or other items that typically appear at the beginning of a new turn, e.g., “iya (no)” and “soo-desu-*ne* (let me see).”
2. We automatically extracted two consecutive utterances produced by two different speakers under the condition that the second utterance started no earlier than the beginning of the last phrase of the first utterance and no later than one second after the end of the first utterance.
3. We only retained utterance pairs where the first utterance initiated an exchange and the second utterance responded to it, such as question-answer pairs, and where a discourse entity overtly expressed in the first utterance was mentioned in the second utterance in the form of either an overt NP or a zero-pronoun. We disregarded the distinction between full NPs and pronouns, both regarding overt NPs. We excluded expressions referring to the interlocutors and overt NPs in the second utterance marked by non-topic particles such as “ga (NOM)” and “ni (DAT).”

These procedures left us 123 instances of utterance pairs: 46 overt and 77 zero expressions in the responding utterance (6 vs. 13 for APS-Int, 20 vs. 26 for SPS-Int, and 20 vs. 38 for Task). Due to the small number of cases in APS-Int, we merged the two interview subsets into a single session type.¹

We next annotated the structure of the responding utterance:
 i) the preface, in which utterance-initial fillers, response tokens,

and other items mentioned above appear; ii) the topic, which is expressed as an overt NP and is repeated from the initiating utterance; and iii) the body, which forms the substantial content of the utterance (see Figure 1). The preface and the topic can be empty; an empty topic means that the discourse entity repeated in the responding utterance is expressed by a zero-pronoun.

2.3. Variables and predictions

The major factor of the analysis was the expression type of the topic in the responding utterance, overt NP vs. zero-pronoun. We examined its effect on two dependent variables. The first dependent variable was the duration of the body. If anti-zero-pronominalization helps the speaker gain time for planning the following utterance, the utterance body would be more complex when the topic is expressed as an overt NP than it is expressed as a zero-pronoun. The complexity of the body is simply approximated by its duration.

The second dependent variable was the presence or absence of the preface. If anti-zero-pronominalization has a similar function to fillers and other prefaced items, overt topics would show a complementary distribution to these items. That is, preface would be less often used when the topic is expressed as an overt NP than it is expressed as a zero-pronoun.

We also investigated whether there is a correlation between the duration of the body and the duration of the topic when the topic is overtly expressed. Since the speaker can gain more time for speech planning when the overt topic becomes longer, the duration of the topic may be positively correlated with the complexity of the body.

All duration variables were log-transformed and standardized before statistical analysis.

2.4. Statistical analysis

To test the above predictions, we employed linear mixed-effects models (Gaussian and logistic models depending on the types of the dependent variables) with random intercept for speakers. We included the session type, i.e., Interview and Task, as a fixed effect. We used lme4 and languageR packages of the R language for model fitting and calculation of MCMC *p*-values [17].

3. Results

3.1 Duration of the body vs. expression type

Figure 2 shows the distribution of the duration of the body depending on the expression type and the session type. The statistical results revealed significant effects of both the expression type and the session type (Table 1); the duration of the body was significantly longer for overt topics than for zero topics and significantly shorter for task-oriented dialogs than interviews.

¹ Keeping these two subsets as independent session types did not affect the conclusions obtained in this paper; the tendency of each session was the same.

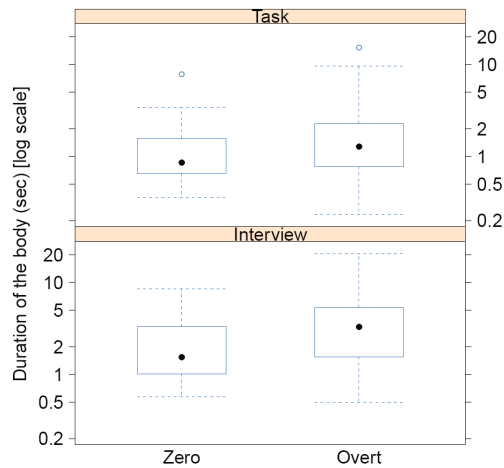


Figure 2: Duration of the body relative to the expression type

Table 1: Estimated parameters: Duration of the body vs. expression type. σ_S indicates the standard deviation at the speaker level and σ indicates the residual standard deviation.

	Coef.	SE	t	MCMC p
(Intercept)	.191	.151	1.268	.237
ExpType:Overt	.426	.171	2.499	.016
SesType:Task	-.738	.181	-4.087	.000

$\sigma_S = .171, \sigma = .907$

3.2. Presence of the preface vs. expression type

Figure 3 shows the ratio of the presence (right bars) vs. absence (left bars) of the preface depending on the expression type and the session type. The statistical results revealed only a significant effect of the expression type (Table 2); prefaces were significantly less often used for overt topics than for zero topics.

3.3. Duration of the body vs. duration of the topic

Figure 4 shows the scatter plot between the duration of the body and that of the topic depending on the session type, when the topic is expressed as an overt NP. The statistical results revealed a significant effect of the session type but no significant effect of the duration of the topic (Table 3); there was no reliable correlations between the duration of the body and that of the topic.

When we focus on the last mora of the topic, instead of the whole topic phrase, however, the situation dramatically changes. Figure 5 shows the scatter plot between the duration of the body and that of the *last mora* of the topic depending on the session type, when the topic is expressed as an overt NP. The statistical results revealed a significant effect of the duration of the last mora in addition to a significant effect of the session type (Table 4); the duration of the body became longer when the duration of the last mora of the topic became longer.

4. Discussion

So far we have shown the following.²

² In addition to these findings, we found that the duration of the body was consistently shorter in task-oriented dialogs than in interviews. This is natural considering that responses by interviewees are generally complex and tend to be long.

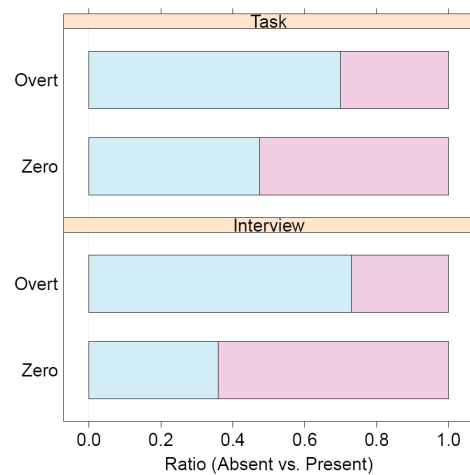


Figure 3: Presence of the preface relative to the expression type

Table 2: Estimated parameters: Presence of the preface vs. expression type. Note that this is a logistic ANOVA model.

	Coef.	SE	z	p
(Intercept)	.475	.301	1.577	.115
ExpType:Overt	-1.294	.403	-3.207	.001
SesType:Task	-.268	.381	-.704	.481

$\sigma_S = .000$

1. The duration of the body is longer when the topic is expressed as an overt NP than it is expressed as a zero-pronoun.
2. Preface is less often used when the topic is expressed as an overt NP than it is expressed as a zero-pronoun.
3. The duration of the body has no correlation with the duration of the whole topic phrase but has positive correlation with the duration of the *last mora* of the topic.

The first result suggests that anti-zero-pronominalization of the topic is a phenomenon closely related to the complexity of the utterance. When the speaker experiences a heavy cognitive load in producing an utterance, s/he tends to use an overt NP, instead of a zero-pronoun, to gain time for planning. In this respect, anti-zero-pronominalization has a similar function to fillers and other speech disfluencies.

The second result provides further support for this theory. Overt topics show a complementary distribution to fillers and other prefaced items, suggesting that the former can substitute for the latter, which has time-gaining function among others.

More interestingly, as shown in the third result, prolongation of the overt topic is enlarged when the substantial content of the utterance becomes complex. Watanabe and Den [13] found that the duration of the last mora of *wa*-marked topic phrases becomes longer as the complexity of the following clause increases, suggesting that the topic marker *wa* would be among time-gaining items. Our result is consistent with this, although about a quarter of the overt topics in our data did not end with *wa* or even lacked an overt topic marker. Thus, their result may be generalizable to topic phrases in general, not limited to *wa*-marked ones.

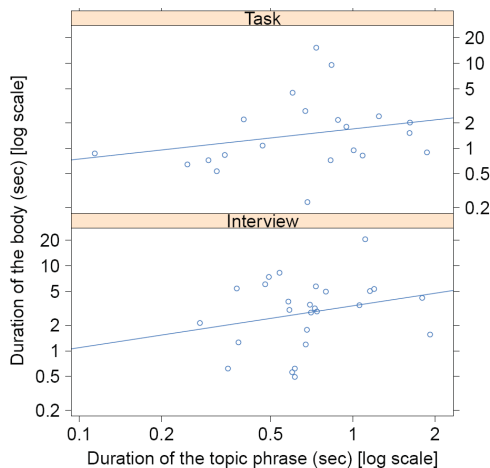


Figure 4: Duration of the body as function of the duration of the topic. Regression lines are also shown.

Table 3. Estimated parameters: Duration of the body vs. duration of the topic.

	Coef.	SE	<i>t</i>	MCMC <i>p</i>
(Intercept)	.360	.242	1.486	.146
DurTopic	.223	.136	1.636	.109
SesType:Task	-.924	.341	-2.710	.015

$\sigma_S = .372, \sigma = .879$

As a line of studies have shown, Japanese has several options to deal with cognitive load at the beginning of an utterance, i.e., fillers [10], word repetitions [12], prolongation [13], and anti-zero-pronominalization (this paper). However, we do not know yet how speakers choose one, or more, option among these on a particular occasion. Investigation into factors behind this choice would be another step in speech disfluency research.

5. Acknowledgements

This work is supported by Grant-in-Aid for Scientific Research (B) “Constructing cognitive and communicative models for dialog based on utterance-unit annotations” led by Yasuharu Den.

6. References

[1] S. Kuno, *The structure of the Japanese language*. Cambridge, MA: MIT Press, 1973.

[2] P. M. Clancy, “Referential choice in English and Japanese narrative discourse,” in *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, W. Chafe, Ed. Norwood: Ablex, 1980, pp. 127–202.

[3] J. Hinds, “Topic continuity in Japanese,” in *Topic continuity in discourse*, T. Givón, Ed. Amsterdam: John Benjamins, 1983, pp. 43–93.

[4] M. Kameyama, “Zero anaphora: The case of Japanese,” Ph.D. dissertation, Stanford University, 1985.

[5] M. Walker, M. Iida, and S. Cote, “Japanese discourse and the process of centering,” *Computational Linguistics*, vol. 20, no. 2, pp. 193–232, 1994.

[6] J. Fry, *Ellipsis and wa-marking in Japanese conversation*. New York: Routledge, 2003.

[7] T. Givón, Ed., *Topic continuity in Discourse*. Amsterdam: John Benjamins, 1983.

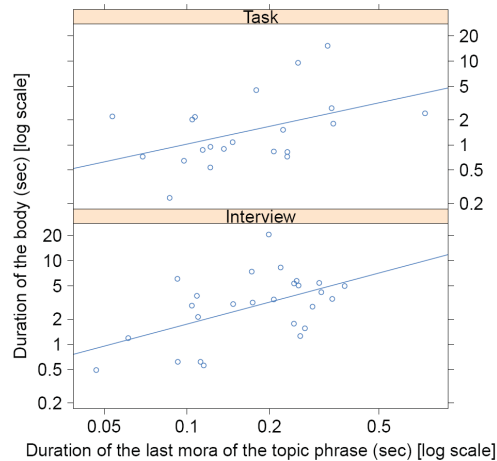


Figure 5: Duration of the body as function of the duration of the last mora of the topic.

Table 4. Estimated parameters: Duration of the body vs. duration of the last mora of the topic.

	Coef.	SE	<i>t</i>	MCMC <i>p</i>
(Intercept)	.268	.166	1.610	.144
DurLastMora	.459	.127	3.626	.001
SesType:Task	-.616	.252	-2.442	.019

$\sigma_S = .000, \sigma = .848$

[8] J. K. Gundel, N. Hedberg, and R. Zacharski, “Cognitive status and the form of referring expressions in discourse,” *Language*, vol. 69, no. 2, pp. 274–307, 1993.

[9] G. Beattie, “Planning units in spontaneous speech: Some evidence from hesitation in speech and speaker gaze direction in conversation,” *Linguistics*, vol. 17, pp. 61–78, 1979.

[10] M. Watanabe, *Features and roles of filled pauses in speech communication: A corpus-based study of spontaneous speech*. Tokyo: Hituzi Syobo, 2009.

[11] H. H. Clark and T. Wasow, “Repeating words in spontaneous speech,” *Cognitive Psychology*, vol. 37, pp. 201–242, 1998.

[12] Y. Den, “Are word repetitions really intended by the speaker?” in *Proc. of the ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech*, Edinburgh, UK, 2001, pp. 25–28.

[13] M. Watanabe and Y. Den, “Utterance-initial elements in Japanese: A comparison among fillers, conjunctions, and topic phrases,” in *Proc. of the DiSS-LPSS Joint Workshop 2010*, Tokyo, 2010, pp. 31–34.

[14] K. Maekawa, “Corpus of Spontaneous Japanese: Its design and evaluation,” in *Proc. of the ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition*, Tokyo, 2003, pp. 7–12.

[15] H. Koiso, Y. Den, and K. Maekawa, “Nihongo-hanasikotoba-koopasu-RDB-no kootiku (Construction of relational database for the Corpus of Spontaneous Japanese),” in *Proc. of the 1st Workshop on Japanese Corpus Linguistics*, Tokyo, 2012, pp. 393–400.

[16] K. Takanashi, T. Maruyama, K. Uchimoto, and H. Isahara, “Identification of “sentences” in spontaneous Japanese: Detection and modification of clause boundaries,” in *Proc. of the ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition*, Tokyo, 2003, pp. 183–186.

[17] R. H. Baayen, *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press, 2008.