

Structure of Perspectivity: A Case of Japanese Reflexive Pronoun “*zibun*”

Yasuhiro Katagiri

NTT Basic Research Laboratories

3-9-11 Midori-cho, Musashino, Tokyo, 180 Japan

katagiri%atom.ntt.jp@relay.cs.net

Abstract

A theory of perspectivity is proposed to establish the foundation of the theory of situated agents. An account is then given, based on the theory of perspectivity, of the use of a seemingly perspectivity related expression, Japanese long-range reflexive “*zibun*.” The theory we proposed for perspectival mental states incorporates two independent notions, *indexicality* and *world view*. The first captures the situatedness of agents within physical environments, and the second captures the mode of reasoning adopted by agents in interacting with other agents. The relationship between these two notions were also discussed. Based on the proposed theory of perspectivity, we argued that, contrary to wide-spread beliefs, the usage of “*zibun*” is not directly related to perspectivity. We gave an alternative explication for the interaction of the usage of “*zibun*” with perspectivity sensitive expressions and the indexical pronoun “*watashi* (I),” in terms of the coreference rule for “*zibun*,” the constraint on the two components of perspectivity, and the agent awareness default principle for the world view.

Introduction

The aim of this paper is twofold. The first is to propose a theory of perspectivity that is reasonable and powerful enough to be the basis of the theory of situated agents. And the second is to give, based on the theory of perspectivity, an account of the use of a seemingly perspectivity related expression, namely, Japanese long-range reflexive “*zibun*.”

Intelligent agents acting in an environment are not outside observers of what’s happening in the world. They are embedded in and interacting with their surrounding environments. For that reason, they perceive, recognize and even describe facts of the environment not from a detached, god’s eye view, but from where they stand in it. Perspectivity, I take, has its origin in this kind of situatedness of agents within environments.

On the other hand, we could and do sometimes imagine a situation from other person’s perspective. How things would look like if I were her? This mode of recognition and reasoning is a fairly common practice

in our daily activities, including such cases as when we are thinking of, trying to explain, or making sense of other person’s behaviors, be they real or imaginary.

Since perspectivity is deeply rooted in human nature as situated agents, it is quite natural that systems of our languages be prepared with mechanisms to reflect and express perspectivity. Different from English reflexive pronouns, which should be co-referential with the subject noun phrases of the same clauses that reflexives appear in, there are several languages including Japanese that have so called long-range reflexives, which could have their antecedents outside the clauses where original reflexives appear. The use of these long-range reflexives have been accounted for by several authors in terms of such notions as empathy[Kuno & Kaburaki, 1977, Kuno, 1987] and logophoricity[Kameyama, 1984, Sells, 1987, Iida & Sells, 1988]. Simply put, all of these accounts are based on the intuition that the use of these long-range reflexives are governed by perspectivity, i.e., from whose point of view the sentence is uttered.

I would argue in this paper first that perspectivity phenomena have two independent ingredients, which I would call *indexicality* and *world view*, and second that, contrary to the explications currently wide-spread, the use of Japanese reflexive “*zibun*” is not directly related to perspectivity. I will give an alternative explanation for the apparent interaction of “*zibun*” and perspectivity based on the two-part theory of perspectivity.

Perspectivity and perspectival utterances

Our perspectival recognition of environments must primarily be reflected in the structure and organization of our representations. Based on our perspectival representation of the environment, when we use language, we sometimes use expressions that are sensitive to perspectivity. So, perspectivity in our representation must be reflected in both the system and the use of our language.

The following characteristics is particularly important when we think of perspectivity of our representation.

- **Missing arguments**

An agent's perspectival recognition of facts amounts to the omission of certain arguments for relations in them. To put it differently, in a perspectival representation an agent is adopting argument-reduced relations.

Think of the relation LEFT-OF. My coffee cup being left of my computer terminal is actually a fact consisting of the three-place relation LEFT-OF₃, my coffee cup, my computer terminal, and myself; my coffee cup is left of my computer terminal seen from me. But I do not and need not be aware of the third argument, myself, when I move my left hand to reach for the cup. This argument is missing in my representation. To say it differently, I am, in that case, conceiving of this fact as if it consisted of the two-place relation LEFT-OF₂, my coffee cup, and the terminal.

In addition to taking into account missing arguments, giving reasonable accounts to the following two phenomena has also to be taken as the requirement for any theory of perspectivity.

- **Account of transferability**

Not only we can recognize our environment from points of view of ourselves, we can take on others' perspectives and could conceive of what the environment would look like if we were they. This ability of assuming other agent's perspective is one of the central functions that differentiates perspectivity from self-identity. The goal of perspective transfer is normally restricted to animate agents.

- **Account of interaction between representation and language**

In addition to the fact that in perspectival utterances the person whose perspective the speaker is assuming is often omitted and made implicit in sentences uttered, each language has a set of perspectivity sensitive expressions. The theory of perspectivity has to give an account of the interaction between perspectival representation of the speaker and the linguistic expressions used in utterances.

Japanese long-range reflexive "zibun"

Each language has its own repertoire of perspectivity sensitive expressions, and Japanese, too, has its own distinctive set of them. Moreover, Japanese long-range reflexive "zibun" has been claimed to interact with them [Kuno & Kaburaki, 1977, Kameyama, 1984]. In (1), lexical item "iku (go)"¹ requires the speaker's point of view to be placed on the person at the source of the movement, which is realized as the subject of the sentence, e.g., "Taro." Since the speaker herself has to be given the highest priority in the set of people on which to place the speaker's point of view, the sentence (2), which requires a speaker to place her point

of view on Taro rather than on herself, is judged bad for most Japanese native speakers.

- (1) Taro-ga Hanako-no ie-ni itta.
 SBJ GEN house-GOAL went
 (Taro went to Hanako's house.)
- ?(2) Taro-ga watashi-no ie-ni itta.
 SBJ I-GEN house-GOAL went
 (Taro went to my house.)

Similarly, the auxiliary verb "kureru"² indicates that the speaker's point of view is placed on the person at the indirect object (IOBJ) position. Due to this restriction posed by "kureru," while (3) is acceptable with the speaker's point of view being placed on the indirect object "Hanako," (4) is bad because of the conflict in perspectivity requirement of "kureru" and that of "watashi (I)."

- (3) Taro-ga Hanako-ni hon-wo
 SBJ IOBJ book-OBJ
 yonde-kureta.
 read -POV-IOBJ
 (Taro read a book for Hanako.)
- * (4) Watashi-ga Hanako-ni hon-wo
 I -SBJ IOBJ book-OBJ
 yonde-kureta.
 read -POV-IOBJ
 (I read a book for Hanako.)

Distribution of "zibun" is apparently correlated with the use of these perspectivity sensitive expressions. Comparable to the contrast between (1)(3) and (2)(4), the sentence (5) is acceptable, whereas (6) is not.

- (5) Hanako-wa Taro-ga zibun-ni hon-wo
 TOP SBJ SELF-IOBJ book-OBJ
 yonde-kureta koto-wo oboeteiru.
 read -POV-IOBJ NOM-OBJ remember
 (Hanako_i remembered that Taro read a book for self_i.)
- * (6) Hanako-wa zibun-ga Taro-ni hon-wo
 TOP SELF-SBJ IOBJ book-OBJ
 yonde-kureta koto-wo oboeteiru.
 read -POV-IOBJ NOM-OBJ remember
 (Hanako_i remembered that self_i read a book for Taro.)

Both in (5) and (6), the use of the auxiliary verb "kureru" requires the speaker's point of view to be placed on indirect objects, that is, "zibun" in (5), but Taro and not "zibun" in (6). So, if we assume "zibun" refers to the person whose perspective the speaker is taking on, these acceptability judgments of (5) and (6) could well be explained.

These observations apparently suggest that the function of "zibun" is to refer to the point of view location.

¹ "itta" is the agglutination of the root of the verb "iku (go)" and the past tense marker "ta."

² "kureta" is the agglutination of the root of the auxiliary verb "kureru" and the past tense marker "ta."

But, the sentences below are known to be perfectly acceptable Japanese sentences[Iida & Sells, 1988].

- (7) Taro-wa zibun-ga kaita hon-wo
 TOP SELF-SBJ write book-OBJ
 watashi-ni yonde-kureta.
 I -IOBJ read-POV-IOBJ
 (Taro_i read the book self_i wrote to me.)
- (8) Taro-wa watashi-ga zibun-wo buuta
 TOP I-SBJ SELF-OBJ hit
 koto-wo oboeteiru.
 NOM -OBJ remember
 (Taro_i remembers that I hit self_i.)

Both in (7) and (8), “*zibun*” is acceptable without referring to the speaker. Furthermore, in (7), point of view location do not even coincide with “*zibun*”. The use of “*kureru*” indicates that the speaker is uttering the sentence from her own perspective, whereas “*zibun*” refers successfully to Taro.

These sentences apparently suggest, contrary to (5) and (6), that the use of “*zibun*” may not have any straightforward relationship to the notion of perspectivity. Any account of the use of “*zibun*” has to give a reasonable explanation to these conflicting intuitions for the relationship between perspectivity and “*zibun*”.

Structure of perspectivity

Apparent conflict concerning the relationship between usages of “*zibun*” and perspectivity described above suggests that the notion of perspectivity needs some clarification. I will propose below a theory of perspectivity which claims that perspectivity actually consists of two independent notions, which I call *indexicality* and *world view*.

Indexicality

The first constituent of perspectivity is related to the situatedness of agents. An agent representing a certain state in her surrounding environment is in a mental state that is in some sense structurally isomorphic to the things represented. Under a situation semantical framework[Barwise & Perry, 1983], an agent A’s believing that a block U is on another block V is a situation classified by the following state of affairs(SOA).

$$\langle\langle \text{Bel}, A, [\dot{s}|\dot{s} \models \langle\langle \text{on}, \dot{x}, \dot{y} \rangle\rangle] \rangle\rangle \\ \wedge \langle\langle \text{of}, \dot{x}, U \rangle\rangle \wedge \langle\langle \text{of}, \dot{y}, V \rangle\rangle$$

A’s mental state itself is classified by a parametric type of situation, $[\dot{s}|\dot{s} \models \langle\langle \text{on}, \dot{x}, \dot{y} \rangle\rangle]$, which shows that the mental state of A is structurally similar to the state represented, e.g., a situation where the block U is on V. Parameters \dot{x} and \dot{y} correspond respectively to A’s concepts of blocks U and V. The second and the third conjuncts in the above formula show that these two concepts are anchored to the real objects U and V, hence

guaranteeing that A’s mental state is really a representation of her surrounding environment.

We observed that, in a perspectival representation of an environmental state, the entity from whose point of view the state is represented, typically the agent herself, is omitted and functions as a missing argument. This is probably made possible by the fact that the agent is embedded in the environment and plays the particular role of the center of coordinates in her representation. All the perceptual information converges to the agent, and all the actions originate from the agent.

We proposed elsewhere[Katagiri, 1989] to represent this kind of indexicality by incorporating a complex relation that has a reduced number of arguments to classify the agent’s perspectival mental states. Consider the block U located to the left of another block V seen from the agent A. A’s perspectival recognition of this situation can be classified by the SOA below, which incorporates a two-place complex relation $\text{LEFT-OF}_{\text{pov}}$ constructed out of the three-place relation LEFT-OF .

$$\langle\langle \text{Bel}, A, [\dot{s}|\dot{s} \models \langle\langle \text{LEFT-OF}_{\text{pov}}, \dot{x}, \dot{y} \rangle\rangle \wedge \langle\langle =, \text{pov}, \text{self} \rangle\rangle] \rangle\rangle \\ \wedge \langle\langle \text{of}, \dot{x}, U \rangle\rangle \wedge \langle\langle \text{of}, \dot{y}, V \rangle\rangle$$

where the relation $\text{LEFT-OF}_{\text{pov}}$ is a complex relation,
 $[\dot{x}, \dot{y} | \langle\langle \text{LEFT-OF}, \dot{x}, \dot{y}, \text{pov} \rangle\rangle]$.

The SOA $\langle\langle =, \text{pov}, \text{self} \rangle\rangle$ in A’s belief shows that she is taking on her own perspective. The point of this representation is that the use of two-place complex relation $\text{LEFT-OF}_{\text{pov}}$ is not merely for notational convenience, but the relation itself has significance in actual reasoning processes within A’s perspectival mental states.

World view

Under the situation semantical formulation described above, the essential constituent determining an agent’s mental state is the type of situation classifying the mental state, e.g., the second argument of the Bel SOA. And the type of situation in turn is determined by the set of parametric SOAs that are constituents of that type. So, we could regard the set of parametric SOAs as the agent’s mental state representing her surrounding environment. When the set includes a SOA whose major constituent is a perspectival relation, the mental state is itself perspectival.

Mental states could represent not only our physical environments but also our social environments. Among them are other agents and the way other agents conceive of their surrounding environments. As is shown in (5)–(8), “*zibun*” is used mostly in sentences describing other agent’s actions and states. In these sentences, uses of each clauses and descriptions have two possible origins. They could be based either on the judgment of the speaker herself, or on the judgment of the agent.³

³More precisely, on the belief of the speaker about the judgment of the agent.

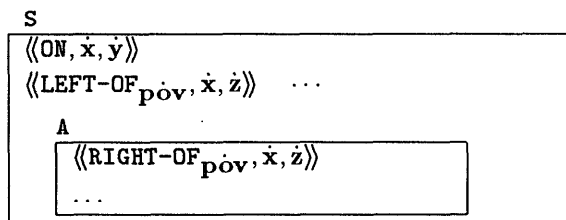


Figure 1: World view of speaker *S* and agent *A*

This division of judgments is depicted in figure 1. The speaker *S* has two sets of SOAs, one for her own conception of the environment, the other for (her conception of) the other agent's (*A*'s) conception of the environment. We call these sets of SOAs *world views*. The set of SOAs at the top level is *S*'s world view, whereas the set embedded is the world view of *A*'s (within *S*). *S* could use either of the world views in reasoning and in constructing expressions to make utterances.

Sentence production process

In a more serious model of language production, we have to take into account the fact that the speaker actually chooses an expression in uttering a sentence based on her conception of what is mutually believed by both the speaker and the hearer [Clark & Carlson, 1982, Appelt, 1985]. But we ignore in this paper this mutual belief aspect involved in language production. This simplification is irrelevant for the essential point made in this paper. We assume here that, in producing sentences, particularly ones describing other agent's actions and states, the speaker chooses, for every description that comprises a noun phrase and for every sub-clause, to adopt either her own world view or the other agent's world view.

Consider an utterance of the following sentence.

- (9) Taro-wa ano koukana ie-wo
 TOP that expensive house-OBJ
 yasui to itta.
 cheap COMP said
 (Taro said that that expensive house is cheap.)

The only consistent interpretation of the utterance is that the description of the house being expensive is within the speaker's world view, whereas the assertion of the house being cheap is within Taro's world view.

Interaction between indexicality and world view

So far, the two dimensions of perspectivity, indexicality and world view, are considered independently. But, there are restrictions on what values we could assume along these two dimensions. The table below summarizes the constraint.

world view(<i>W</i>) indexicality(<i>I</i>)	Speaker(<i>S</i>)	Agent(<i>A</i>)
Speaker(<i>S</i>)	○	×
Agent(<i>A</i>)	○	○

$I(\text{indexicality})=W(\text{world view})=S(\text{speaker})$ means that the speaker is taking her own perspective, whereas $I=W=A(\text{agent})$ corresponds to cases where the speaker is taking on the other agent *A*'s perspective.

The asymmetry between $I=A\&W=S$, which is legitimate, and $I=S\&W=A$, which is not, is because in the former situation, the speaker *S* raises *A*'s *pov* value to her own world view, hence partially taking on *A*'s perspective with her own descriptive framework, while to realize the latter combination, the speaker *S* has to embed her own *pov* value into the world view of the agent *A*, which is impossible, since *S* knows that her own *pov* value is outside the *A*'s embedded world view and hence inaccessible from it.

The former situation, $I=A\&W=S$, is typical in a person having heard and comprehended a perspectival utterance. Upon comprehending what a speaker's perspectival utterance means, the hearer normally takes only on the speaker's *I* value, but adopts her own stock of descriptions to describe/identify objects.

Agent awareness default principle

There is another constraint concerning world view and language production. This constraint is not the one which has to be strictly observed. Rather, it provides a default value for the world view in producing utterances.

Agent awareness default principle

When describing actions and states of an agent, the speaker adopts the agent world view by default.

Agents usually perform actions consciously. They are conscious of the facts that they will perform, are performing, and have performed the actions. And they are aware of facts that had held before, having been held throughout, and were established by the actions. Hence, it is natural to assume, in an utterance reporting an agent's behavior, descriptions, and hence world view, are usually taken to be of the agent's by default.

Perspectivity and "zibun"

Having developed the underlying theory of perspectivity, we can now state the hypothesis concerning the usage of "zibun" in the form of a coreference rule. Note that the rule below is not itself directly related to the notion of perspectivity.

Coreference rule for "zibun"

The use of "zibun" is based on the judgment of identity of the referent of "zibun" to the logical agent of an action or to the logical experiencer of a mental state described in the sentence.⁴

⁴By "logical" I mean here that we take these case roles as relations at the level of semantic representation, which are relatively independent of surface syntactic realizations.

The point of the coreference rule above is that the identity judgment for the use of “*zibun*” could either be based on the speaker’s world view or on the agent’s. Although, by the agent awareness default principle, the judgment is taken to be of the agent’s by default, in a special circumstances where the principle could be defeated, the judgment can be strictly of the speaker’s. The following sentence with “*zibun*” referring to Taro can be used without any problems to describe Taro’s amnesia story. Since Taro, being amnesiac, does not have the knowledge of his having been a baseball player, the agent awareness default has to be defeated and the identity judgment for “*zibun*” in this case is based on the speaker’s world view.

- (10) Taro-wa zibun-ga yakyuu-senshu datta
 TOP SELF-SBJ baseball-player was
 koto-wo shira-nai.
 NOM -OBJ know-NEG
 (Taro_i does not know that self_i was a baseball player.)

Interaction of “*zibun*” and perspectivity

Given the constraint on indexicality and world view, and the coreference rule for “*zibun*,” we can give an explanation to the possibility of apparent conflict between the usage of “*zibun*” and perspectivity phenomena, namely, why sentences (7) and (8) are acceptable.

According to the constraint, when an expression is originated from the speaker’s world view ($W=S$), there are two possibilities for the indexicality value, the speaker and the agent ($I=S$ or $I=A$). If the speaker takes herself (S) as the indexicality value, “*zibun*” could refer to an agent without the agent necessarily occupying the value of indexicality $p\acute{o}v$.

In both of the sentences (7) and (8), we could interpret the judgment of the coreference of “*zibun*” and the agent of reading the book or the experiencer of remembering as originating from the speaker’s world view. Under that interpretation, these sentences can be construed without problems, even if we take the speaker as the indexicality value.

After freeing “*zibun*” from perspectivity, the remaining task is to explain apparent interaction between the two; namely, why (6) is judged bad. An explanation can be found by looking at the agent awareness default principle.

The use of “*oboeteiru* (remember)” in (6) enforces the application of the agent awareness default principle, and the complement clause, what is remembered, is taken by default to be within the agent Hanako’s world view. But once that default value is assumed, the constraint on world view and indexicality restricts the possible value for indexicality to the agent Hanako. On the other hand, the auxiliary verb “*kureru*” requires the indexicality value to be the one in the indirect object position, e.g., Taro. The sentence is judged bad

because of the incoherence caused by these two conflicting requirements.

The subtlety of the interaction of agent awareness default principle and perspectivity could be seen by the following pair of sentences.

- (11) Taro-wa zibun-no tsuma-no migigawani
 TOP SELF-GEN wife-GEN to-the-right
 Hanako-ga iru noni kigatsuita.
 SBJ be COMP noticed
 (Taro_i noticed that Hanako was to the right of self_i’s wife.)
- (12) Taro-wa zibun-no tsuma-no migigawani
 TOP SELF-GEN wife-GEN to-the-right
 Hanako-ga iru noni kigatsuitei-nai youda.
 SBJ was COMP noticed-NEG seem
 (Taro_i seems not to have noticed that Hanako is to the right of self_i’s wife.)

In (11), the use of “*kigatsuita* (notice)” enforces the agent awareness default and by the chain of reasoning similar to the one for (6), indexicality value, which fills the missing argument of $RIGHT-OF_{p\acute{o}v}$, has to be the agent Taro. In contrast, the negation of noticing in (12) together with uncertainty in judgment could prevent the application of the agent awareness default, and consequently the $RIGHT-OF_{p\acute{o}v}$ of (12) can be interpreted from the point of view of the speaker.

Conclusions

We proposed our theory of perspectivity as a basis for the theory of situated agents. Situatedness has two aspects; situatedness in physical environment, and situatedness in social environment. Each of the two aspects has its own repercussions to the organization of representations. An agent can exploit environment as an extension of its representational medium and put as much information as possible into environment rather than into their own representation. Secondly an agent can use its own representation in reasoning about reasoning performed by other agents. Our notions of indexicality and world view together with constraints on the interaction between the two capture both types of situatedness.

Based on the theory of perspectivity, together with coreference rule for “*zibun*,” we showed that the use of Japanese long-range reflexive “*zibun*” is not directly related to perspectivity, and gave an alternative explanation for the interaction of “*zibun*” and perspectivity. We claimed that “*zibun*” simply corefers with the logical agent/experiencer of a sentence, and that behind the apparent interaction between perspectivity and “*zibun*” lies the interaction among this coreference function of “*zibun*,” the constraint on the two components of perspectivity, and our default assumption on the world view speakers adopt in issuing utterances.

One notable point of the account we gave of the use of “*zibun*” is that, although simple, the account is in

the form of a process theory, which explains the usage of "zibun" in terms of the structure of underlying representation and the mechanism of language production. As the account shows, this type of theory has far wider potential, compared with purely syntactic approach conventional in linguistics, for explaining *uses* of linguistic expressions.

Beyond the problem of perspectivity we discussed in this paper, we could expect our theory to be extended, with relatively straightforward enrichment, to provide us with a new way of thinking of the concept of the discourse focus [Sidner, 1983] or that of the center [Grosz *et al.*, 1983, Kameyama, 1986]. Our theory could then be accommodated to phenomena related to the uses of elipsis and anaphora, in general, and to the use of Japanese zero-pronoun, in particular. We are also thinking of applying our theoretical framework to problems in situated planning and plan recognition for situated agents.

Acknowledgment

The author would like to thank Ikumi Imani, Hajime Narita, Akira Shimazu, Ikuo Takeuchi, Syun Tutiya, and Masaaki Yamanashi for their helpful comments, criticisms, and supports.

References

- Appelt, D. E. 1985. *Planning English sentences*. Cambridge University Press.
- Barwise, J. and Perry, J. 1983. *Situations and Attitudes*. MIT Press.
- Clark, H. H. and Carlson, T. B. 1982. Speech acts and hearers' beliefs. In Smith, N. V., ed. 1982, *Mutual Knowledge*. Academic Press. chapter 1, 1-36.
- Grosz, B. J.; Joshi, A. K.; and Weinstein, S. 1983. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st Annual Meeting of the Association of Computational Linguistics*. 44-50.
- Iida, M. and Sells, P. 1988. Discourse factors in the binding of zibun. In Poser, W. J., ed. 1988, *Papers from the Second International Workshop on Japanese Syntax*. Center for the Study of Language and Information, Stanford University. 23-46.
- Kameyama, M. 1984. Subjective/logophoric bound anaphora *zibun*. *Chicago Linguistic Society* 20:228-238.
- Kameyama, M. 1986. A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meeting of the American Association for Computational Linguistics*. 200-206.
- Katagiri, Y. 1989. Semantics of perspectival utterances. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*. Morgan Kaufmann. 1474-1479.
- Kuno, S. and Kaburaki, E. 1977. Empathy and syntax. *Linguistic Inquiry* 8:627-672.
- Kuno, S. 1987. *Functional Syntax: Anaphora, Discourse and Empathy*. The University of Chicago Press.
- Sells, P. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18:445-479.
- Sidner, C. L. 1983. Focusing in the comprehension of definite anaphora. In Brady, M. and Berwick, R. C., eds. 1983, *Computational Models of Discourse*. MIT Press.