

English Relative Constructions and Discourse in Spoken and Written Language

Megumi Okugiri*

Postgraduate School, Tokyo Healthcare University, Tokyo, Japan

(Received 28 November 2013; Final version received 20 February 2014)

This study investigates the frequency of English relative constructions in spoken and written language, explains the different inclinations from the discursal point of view, and analyses the information status of the head and relative clause types. Samples of 735 relative constructions are extracted from spoken and written corpora. The study reveals the frequency of the head noun phrases and RC-type differing in the two different modes of language. The results show that the head types are likely dependent upon the presence or absence of kinetic information, visual referents, and deictic referents, and that the determining factor for the RC-type may be planning time. Through the exhibition of different prototypical patterns of English relative constructions between spoken and written language, discursal properties are shown to play a crucial role in the production of English relative constructions.

本研究は、話し言葉と書き言葉における英語関係節構文の産出頻度を調査し、その結果を談話的に分析するものである。話し言葉コーパスと書き言葉コーパスから合計 735 の英語ネイティブスピーカーによる関係節構文を抽出し、関係節構文の先行詞と関係節のタイプに着目した。その結果、話し言葉と書き言葉では、産出が違うことがわかった。特に先行詞の違いについては、目視による情報、また目視できる直示的な対象物の有無が、そして関係節のタイプについては、プランニング時間の有無が、話し言葉と書き言葉での関係節構文の違いにおける主要な要因であると推測される。以上の結果から、話し言葉と書き言葉では、関係節をはじめとした構文の産出において違いが見られ、談話内での様々な要因がその産出に影響していると提言する。

1 Introduction

The aim of this study is to investigate the tendency towards English relative constructions in spoken and written language and explain the different inclinations from the discursal point of view. This paper will also analyse the information status of the head and relative clause types in order to reveal discursal factors affecting relative clause production.

The Noun Phrase Accessibility Hierarchy is a well-known typological hierarchy predicting the presence of relative clause (RC) types in natural languages in terms of markedness (Comrie & Keenan, 1979). However, this hierarchy does not explain the acquisition of semantic or discursal properties by relative constructions. Other studies,

*Email: m-okugiri@thcu.ac.jp

such as those of Kidd, Brandt, Lieven, and Tomasello (2007), Mak, Vonk, and Schriefers (2006), Traxler, Morris, and Seely (2002) and Diessel and Tomasello (2000, 2005) find slightly different results. These studies focus on the acquisition of relative constructions in terms of semantic properties such as animacy, topic-worthiness, and semantic complexity, and find subject RCs to be strongly associated with animate head noun phrases (NPs) and object RCs with inanimate head NPs.

In terms of RC functions, Fox and Thompson (1990) investigated the use of RCs in adult English conversations, observing their use in identifying and characterising head noun phrases. Fox and Thompson also revealed the relationship between functions and RC-types, highlighting the tendency of subject RCs to *characterise* the head NPs, while object RCs tended to *identify* the head NPs; these tendencies were associated with preferences in regard to argument type (Du Bois 1987, 2003). A strong preference emerges for transitive subject NPs to be Given referents (*the Non-Lexical A Constraint*), typically pronouns.

In the current study, 735 relative constructions are extracted from spoken and written corpora. It will reveal that the frequency of the head and RC-types were divergent in the two different modes of language. The results show that the head types depend on the presence or absence of kinetic information, visual referents, and deictic referents, and that the major factor in determining RC-type may be planning time. It will also present different prototypical patterns of English relative constructions between spoken and written language.

2 Previous Studies

In previous studies of relative construction, much of the focus has been placed on the accordance of the acquisition order using the NPAH, a typological hypothesis proposed by Comrie and Keenan (1979) to predict the presence of RCs in natural languages. Their hypothesis predicts the presence of each RC in the order of subject > direct object > indirect object > oblique > genitive > object of comparative particle.

The NPAH, however, predicts only the syntactic properties within an RC. It does not predict either the acquisition of the heads modified by RCs nor the semantic and discoursal properties of relative constructions. One of the studies focusing on the semantic properties of RCs was Kidd et al. (2007), which demonstrated the importance of semantic properties, revealing that English and German children, like adults, produce object relatives when the head noun is inanimate and when the subject of the RC is pronominal in both languages. Thus, the study showed that subject RC is not always a child's first construction as the NPAH predicts. Traxler et al. (2002), Mak et al. (2006), and Diessel (2009) found similar results, arguing the existence of a strong association between agentivity and RC-types. Mak et al. (2006) proposed that the processing difficulty of Dutch RCs in a reading task is explainable by the interaction of the animacy of the subject as a topic and the RC-type: the readers preferred animate entities to be the subject of the RC; i.e. animate heads are highly plausible candidates for agents, thus they result in a subject RC. This result is supported from the discoursal point of view. As Du Bois (1980) and Fox and Thompson (1990) maintain, inanimate referents,

being less agentive, tend to be non-subject NPs, and more agentive animate referents tend to be subject NPs. Furthermore, Fox and Thompson argue that inanimate referents are often made relevant in discourse by relating them to the humans (animate referents) who own and use them, which are highly agentive in L1 English general discourse. Ming and Chen (2010) demonstrated that animacy also plays an important role in the production of Chinese relative clauses.

Moreover, Fox and Thompson (1990) found relative constructions in adult English speech to be associated with constraints in regard to argument type (Du Bois, 2002), i.e., *the Non-Lexical A Constraint*. The constraint predicts a strong preference for transitive subject NPs to be given information, typically pronouns. Thus, their study revealed a strong discursal effect on the NP-types of the relative constructions in English conversation.

Regarding NP types in relative constructions, Reali and Christiansen (2007) investigated the frequency and distribution of pronouns and lexical NPs in object and subject RCs within adult English speakers' corpus data. They compared the frequencies at which their speakers encountered difficulty during the online processing of relative constructions in self-paced reading experiments. The authors found the same results as Kidd et al. (2007) in both the corpus study and the online experiments. Reali and Christiansen's (2007) findings are noteworthy insofar as they demonstrate that the potential difficulties in comprehension encountered during online processing mirror the frequency of distributional patterns in language. Diessel's (2008, 2009) finding supports those of Reali and Christiansen (2007).

Diessel and Tomasello (2000) discuss the function of children's relative constructions in English and reveal that the production was greatly affected by semantic complexity. They examined the speech of four children at different stages of acquisition and found that the children produced presentational relative constructions at an early stage of acquisition. These constructions include a predicate NP as the main clause with a subject RC and the relative pronoun (Diessel & Tomasello, 2000, p. 135):

(1) Is this something that turn around?¹

In (1), the main clause is a copular sentence and is semantically empty. Therefore, a presentational relative construction, like an amalgam construction, is semantically less complex than other kinds of relative constructions as it contains only a single proposition in the whole sentence. Lambrecht (1988) explains this as the presentational amalgam construction that is frequently found in adult English speech. After acquiring these types of relative constructions, the children produced more complex relative constructions. Thus, Diessel and Tomasello argued that children's acquisition of relative constructions is affected by semantic complexity (Diessel & Tomasello, 2005; Diessel, 2009). Fitz and Chang (2008) found a similar result in their connectionist experiment in which a computational model had to learn the various types of English RCs from a

¹ The expected sentence is "Is this something that turns around?"

training sample of simple and complex sentences.

The tendencies found in these studies are associated with preferences in regard to argument type, termed the Preferred Argument Structure as proposed by Du Bois (1987, 2002, 2003). Du Bois (2002) examined English native speakers' speech data in the Santa Barbara Corpus of Spoken American English (Du Bois, 2000) and analysed all NPs not limited to relative constructions. He argues that the argument type is triggered by verbs and proposed a constraint formulated from the frequency of the pattern of arguments in discourse—*the Non-Lexical A Constraint*²—which postulates a preference for avoiding a lexical NP in the transitive subject position. That is, for NPs in the intransitive subject, object and oblique position, lexical nouns occur freely, while for NPs in the transitive subject position, pronouns are used far more frequently, as in (2).

(2) *I* always send *letters* to *my mother*.

Du Bois' analysis supports the idea that speakers' or acquirers' language exhibits a discursal preference for certain syntactic forms.

Thompson (1997) clarified the degree of lexical transitive subject NP, as well as the information status of all kinds of NPs in English in adult conversation. He found a tendency similar to that of Du Bois (2002), and determined that NPs in the oblique position were most likely to be New. A New referent is newly introduced into the discourse and presumed not to be in the hearer/reader's focal consciousness or understanding. The percentages of New NPs in each argument position were 65% in the oblique position, 35% in the object and intransitive position, and only 15% in the transitive subject position. The percentages of New NPs in Thompson's (1997) study are summarised in Table 1.

Table 1. Percentages of New/Given Noun Phrases in English

	English	
	New	Given
A (transitive subject)	15 %	85 %
S (intransitive subject)	35 %	65 %
O (object)	35 %	65 %
B (oblique)	65 %	35 %

A Given referent is presumed to be in the hearer's focal consciousness and is distinguished as a referent that is anaphorically linked to its previous mention. Thompson found 85% of transitive subject NPs to be Given referents, typically represented with lexical NPs.

Chafe (1980) and Du Bois (1980) proposed another type of information status: an Identifiable referent that is presumed not to be the hearer's present focal consciousness,

² Here, "A" is referred to as a transitive subject.

but for which the hearer is able to identify the referent as having a strong correlation with a previously-introduced referent in terms of frame, prior discourse, or prior knowledge (including world knowledge.) Since Thompson (1997) did not include the Identifiable referent in his study, the current study will examine all three types of information status and attempt to find the preferred information status of head NPs modified by relative constructions in order to find the effect of discourse. To avoid ambiguity with other uses of these words in English, the three terms are capitalised when used to indicate the information status of the heads.

This study focuses on production data and assumes that frequency is a key determinant in discovering prototypical language use. Frequency reveals the nature of language from the usage-based approach (Diessel, 2004; Tomasello, 2003), in which utterances are defined as strings of speech used to make things understood, and where the strings constitute a construction that has meaning (Lieven & Tomasello, 2008; Tomasello, 2003). Furthermore, Bybee (2008) suggests that the most frequent strings of speech have the strongest representations in memory and serve as analogical bases for forming novel instances of the construction; the most frequent string is the central, prototypical pattern. Thus, the present study seeks to find the prototype of English relative constructions by analyzing frequency.

One may reasonably assume the frequency of a linguistic item to be different between spoken and written language modes. Brown and Yule (1983) argue that spoken and written language fulfill different primary functions. The primary function of general written language is transactional, relating to the transmission of information, whereas that of general spoken language is also interactional, relating to establishing or maintaining a social relationship.

Chafe (1984a) more explicitly differentiates spoken and written language when he argues that speaking and writing are completely different activities. Spoken language tends to be linguistically fragmented, where simpler and less information is contained in an information unit. Written language tends to be integrated, where more complex and denser information is present. Chafe (1984b) argues that formality is also a contributing factor to language forms.

Another difference between spoken and written language involves kinetic and paralinguistic information (Tannen, 1982), visual referents, and deictic referents. In spoken language, kinetic information, such as facial expression and gesture, paralinguistic information, such as tone of voice, intonation, prosody, and pauses (Ochs, 1979), and visual referents and deictic referents are generally available because the addressee is in front of the addresser. In written language, the writer, in the absence of the reader, must anticipate their comprehension (Bygate, 1987). Thus, the writer need only establish cohesion through lexicalisation and complex syntactic structures (Tannen, 1982).

Planning time is another difference between spoken and written modes. Although there are exceptions such as planned speech (e.g., the president's speech) and unplanned writing (e.g., a casual letter to a friend), spoken language generally allows for less planning than written language (Bygate, 1987; Tannen, 1982). Ochs (1979) describes

the difference between unplanned and planned discourse, defining unplanned discourse as “discourse that lacks forethought and organizational preparation” and planned discourse as “discourse that has been thought out and organized (designed) prior to its expression” (Ochs, 1979, p.55). Skehan (1998) and Ochs (1979) argue that planned discourse exhibits more linguistic complexity. The greater the planning, the greater the grammatical complexity.

The present study focuses on the divergence of English relative constructions between spoken and written language. It attempts to find the factors affecting the production in the two modes and reveals that discursual properties play a crucial role in the cognition of language users.

3 Research Questions and Method

Based on the previous studies, the research questions are as follows:

1. Is the prototypical type of English relative construction different between spoken and written language?
2. If there is a difference between spoken and written language, what are the major determining factors?

The relative constructions of native English speakers were extracted from the National Institute of Information and Communications Technology Japanese Learner English Corpus (Izumi et al., 2005) and the Nagoya Interlanguage Corpus of English (Sugiura, 2008)³, the former being a spoken corpus, the latter a written one. Paralinguistic information, such as tone of voice, intonation, prosody, and pauses will not be in consideration here as it is not available in the corpus data. However, the availability of kinetic information, visual referents, and deictic referents is highly possible in the spoken mode. Planning time is available in the written data as well as in the context of data collection⁴. Samples of relative constructions were extracted from 20 files from the spoken corpus and 28 files from the written corpus.

Examples of relative construction were extracted by hand. After extraction, the heads were categorised as intransitive subject (S), transitive subject (A), object (O), oblique (B), noun phrase (NP), predicate NP (PN) and main clause (M). The RCs were also categorised as intransitive subject (SR), transitive subject (AR), object (OR), and

³ The data comes from a part of my PhD dissertation at the University of Tokyo, where I compared the native English speakers' and Japanese learners' language.

⁴ For the spoken corpus, the data collection method consisted of an oral interview test, 15 minutes in length, divided into five parts: a warm-up, a picture description task, a role-play, storytelling and a wind-down. In the 3-4 minute warm-up, the interviewer asks several questions in order to establish a comfortable atmosphere. In the 2-3 minute picture description task, the applicant is required to describe the picture that the interviewer shows them. In the 1-4 minute role-play task, the applicant is told to play a role that the interviewer assigns to him/her. In the 2-3 minutes of storytelling, the applicant is shown four- or six-frame pictures and told to create a story. Finally, in the 1-2 minutes of wind-down, the interviewer asks a last series of questions to calm the applicant.

oblique (BR). Examples are shown in the Appendix. The information status of the head NPs were categorised into New, Given, or Identifiable.

4 Results and Discussion

In total, 735 relative constructions were found in the corpora, 527 in the spoken and 208⁵ in the written. As illustrated in Table 2, the most frequent heads in the spoken mode were PN followed by O then B heads, while in the written mode B was the most frequent followed by PN then S.

Table 2. *Frequency of Heads in Each Mode*

	S	A	O	B	M	NP	PN	total
spoken	42 7.97%	7 1.33%	131 24.86%	122 23.15%	47 8.92%	39 7.40%	139 26.38%	527 100%
written	42 20.19%	10 4.81%	36 17.31%	70 33.65%	3 1.44%	2 0.96%	45 21.63%	208 100%
total	84	17	167	192	50	41	184	735

S=intransitive subject, A=transitive subject, O=object, B=oblique, M=main-clause, NP=noun phrase, PN=predicate noun phrase

A chi-square analysis was performed to examine the interaction between head type and mode. It yielded a significant difference between the head and mode: χ^2 (6, N=735)=62.019, $p<.0001^*$. Therefore, the results revealed different head types depending on the mode and the heads. The types of more frequent heads seem to be consistent with the New referents, namely oblique and object, found in Thompson (1997) except for the S head in the written mode.

The frequency of PN heads is in accordance with the findings by Diessel and Tomasello (2000) and Lambrecht (1988). Moreover, the results of the information status of PN heads in the spoken mode support their findings that the presentational relative construction is used to introduce New referents into discourse. The proportion of New NP in PN heads is 60.43% in the spoken mode as shown in Table 2. This result implies that the pragmatic function to introduce New referents plays an important role in the spoken mode.

⁵ The frequency of the relative constructions in the written mode was 8.73 per 1,000 words. The author was not able to count the frequency for the spoken mode because the available software with the spoken corpus does not count the token and the files include a tremendous number of discourse transcription tags such as <SC> for self correction and <ncs> for non-verbal sound. It is left for my future study to reveal the frequency in order to compare the production in two different modes and to find out how planning time affects the production.

Table 3. *Frequency of New, Given and Identifiable NPs in Each Mode*

Head		New	Given	Identifiable	Total
S	spoken	16 (38.10%)	9 (21.43%)	17 (40.48%)	42
	written	16 (38.10%)	12 (28.57%)	14 (33.33%)	42
A	spoken	4 (57.14%)	0 (0.00%)	3 (42.86%)	7
	written	2 (20.00%)	2 (20.00%)	6 (60.00%)	10
O	spoken	85 (64.89%)	13 (9.92%)	33 (25.19%)	131
	written	20 (55.56%)	8 (22.22%)	8 (22.22%)	36
B	spoken	90 (73.77%)	8 (6.56%)	24 (19.67%)	122
	written	32 (45.71%)	14 (20.00%)	24 (34.29%)	70
M	spoken	35 (74.47%)	1 (2.13%)	11 (23.40%)	47
	written	3 (100.00%)	0 (0.00%)	0 (0.00%)	3
NP	spoken	27 (69.23%)	7 (17.95%)	5 (12.82%)	39
	written	1 (50.00%)	0 (0.00%)	1 (50.00%)	2
PN	spoken	84 (60.43%)	24 (17.27%)	31 (22.30%)	139
	written	22 (48.89%)	12 (26.67%)	11 (24.44%)	45

On the other hand, although the proportion of New PN heads is highest (48.89%) in the results in the written mode, and the statistical results yielded no significant difference (χ^2 (2, N=184)=2.390, $p=0.3027$ *n.s.*), the proportion of Given referents (26.67%) is 9.4% higher than that of spoken. Two such examples in the data are shown in (4).

(4) These are the ones that need the help, these are the ones that have repressed their feeling about what is going on so deep inside that they do not show any signs of problems until they are late into their teens and sometimes later.

The underlined NPs are the heads. They are both Given referents in the example. Due to the absence of kinetic information, visual referents, and deictic referents, the writer probably modified the Given referents with RCs to clarify the referents and maintain consistency in the discourse.

For each head, a chi-square analysis only yielded significant difference for B between the modes (χ^2 (2, N=192)=16.324, $p=.00003^*$). In Table 2, the proportion of B head is 23.15% in the spoken mode. The proportion is 10.5% higher in written mode, at 33.54%. With regard to its information status, the proportion of New B heads in the spoken mode is 73.77%, which Thompson (1997) finds the most likely to be New. Conversely, the proportion of the written mode is much less (45.71%). In addition, the proportions of Given and Identifiable referents in the written mode (20% for Given and 34.29% for the Identifiable) are much higher than that of the spoken mode (6.56% for Given and 19.67% for Identifiable). (5) is an example of an Identifiable B head, while (6) demonstrates a Given B head.

- (5) Many people who commit suicide are stricken with depression or other mental illness that drive them to suicide.
- (6) Another thing that is important that I don't think gets enough attention though are the families and friends involved with the person who dies.

In addition to the complex and denser information typical in written language, the writer and reader share no kinetic information, visual referents, and deictic referents, which serve to burden the delivery of the proposition even more. These may trigger the writer to produce such relative constructions with Given and Identifiable heads more often than they do in spoken mode.

The same seems to be true for the frequent S head in the written mode. In order to maintain coherence, writers need to refer to classes of third-person human referents rather than unavailable pronouns such as *I* or *she* in the written language. Thus, the writers exhibited a greater need to identify referents even though the information is already Given in written language as opposed to spoken language. In fact, the proportions of New S heads are relatively low compared with other head types (in Table 3, 38.10% in both spoken and written).

It is also notable that the proportion of NP is higher in the spoken mode than in the written mode. The result implies that spoken language, as Chafe (1982) argues, is more fragmented.

Regarding RC-type, Table 4 exhibits the frequency and proportions.

Table 4. *Frequency of RCs in Each Mode*

	SR	AR	OR	BR	total
spoken	237 44.97%	78 14.80%	153 29.03%	59 11.20%	527 100%
written	76 36.54%	82 39.42%	26 12.50%	24 11.54%	208 100%
total	313	160	179	83	735

SR=intransitive subject RC, AR=transitive subject RC, OR=object RC, BR=oblique RC

A chi-square analysis yielded a significant difference between the head and mode: $\chi^2(3, N=735)=60.778, p<.0001^*$. Thus, the tendency is different depending on the modes. In the spoken mode SR is the most frequent, while in the written mode AR is the most frequent. It is reasonable to assume that planning time allows writers to produce NPs exhibiting denser and more complex compositions

5 Conclusion

The major findings of this study are that the frequency of English relative construction type is different between spoken and written language, that the factors for the head types seems to be the presence or absence of kinetic information, visual referents, and deictic referents, and the factor for the RC-type may be planning-time. This study demonstrates the great potential of discursal explanation in analysing a grammatical

construction.

However, the study was not able to explain why O head occurred more frequently than B head. One possibility is that a sentence including B head may require an extra NP in addition to the object compared to a sentence with O head as in (7).

- (7) ... they had the chance from the beginning to purposefully shape and adjust their education to model which best fit the societal ideals ...

In (7), the object NP is “their education” and the oblique NP is “model.” The inclusion of two types of NPs, object and oblique, in a sentence is, of course, more complex than a sentence with only one NP. The following sentence (repeated example of (5)) with B head, however, does not include the object.

- (5) Many people who commit suicide are stricken with depression or other mental illness that drive them to suicide.

It will require collecting more samples of RCs from a larger corpus such as British National Corpus and further investigation to find out the prototypical patterns and to draw a definitive conclusion and determine which of and whether kinetic information, visual referents, deictic referents and/or planning time are the definitive factors in relative constructions.

Acknowledgments

This is a part of my PhD dissertation at the University of Tokyo. I am grateful to Professor emeritus Paul Rossiter, Professor Tom Gally, Professor Seiko Fujii, Professor Yasuhiro Shirai, Professor emeritus Hideo Oka, and Professor Yuko Itatsu for their invaluable comments. I also wish to thank anonymous reviewers for their suggestions on an earlier version of this article. I remain solely responsible for any remaining errors.

References

- Brown, G., & Yule, G. (1983). *Teaching the Spoken Language: An Approach Based on the Analysis of Conversation English*. Cambridge, UK: Cambridge University Press.
- Bybee, J. (2008). Usage-based grammar and second language acquisition. In Robinson, P. & Ellis, N. (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, 216-236. New York, NY: Routledge.
- Bygate, M. (1987). *Speaking*. Oxford, UK: Oxford University Press.
- Chafe, W. (1984a). Integration and involvement in spoken and written language. In Borbe, T. (Ed.), *Semiotic Unfolding* (Vol. 2, pp. 1095-1102). Berlin, DE: Mouton de Gruyter.
- Chafe, W. (1984b). Integration and involvement in speaking, writing, and oral literature. In Tannen, D. (Ed.), *Spoken and Written Language: Exploring Orality and Literacy* (pp. 35-53). Norwood, NJ: Ablex.

- Clancy, P. (1980). Referential choice in English and Japanese narrative discourse. In Chafe, W. (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, 127-202. Norwood, NJ: Ablex.
- Comrie, B. & Keenan, E. (1979). Noun phrase accessibility revisited. *Language*, 55(3), 649-664.
- Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge, UK: Cambridge University Press.
- Diessel, H. (2008). The emergence of relative clauses in early child language. In 12th Biennial Rice University Symposium on Language. Retrieved from <http://www.ruf.rice.edu/~eivs/sympo/papers/Diessel.pdf>
- Diessel, H. (2009). On the role of frequency and similarity in the acquisition of subject and non-subject relative clauses. In Givon, T., & Shibatani, M. (Eds.), *Syntactic Complexity* (pp. 251-276). Amsterdam, NL: John Benjamins.
- Diessel, H. & Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics*, 11, 131-151.
- Diessel, H. & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, 81(4), 882-906.
- Du Bois, J. (1980). Beyond definiteness: The trace of identity in discourse. In Chafe, W. (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, 203-274. Norwood, NJ: Ablex.
- Du Bois, J. (1987). The discourse basis of ergativity. *Language*, 63, 805-855.
- Du Bois, J. (2000). *Santa Barbara Corpus of Spoken American English, Part 1. 3 CD-ROMs*. Philadelphia, PA: Linguistic Data Consortium.
- Du Bois, J. (2002). Discourse and grammar. In Tomasello, M. (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Vol. 2, 47-88. Mahwah, NJ: Lawrence Erlbaum.
- Du Bois, J. (2003). Argument structure: Grammar in use. In Du Bois, J., Kumpf, L., & Ashby, W. (eds.), *Preferred Argument Structure: Grammar as Architecture for Function*, 11-60. Amsterdam, NL: John Benjamins Publishing Company.
- Fox, B., & Thompson, S. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language*, 66(2), 297-316.
- Izumi, E., Uchimoto, K. & Isahara, H.. (2005). *The National Institute of Information and Communications Technology Japanese Learner English Corpus*. Tokyo, JP: ALC.
- Kidd, E., Brandt S., Lieven, E., & Tomasello M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, 22(6), 860-897.
- Lambrech, K. (1988). There was a farmer had a dog: Syntactic amalgam revisited. *Proceedings of the annual meeting of the Berkeley Linguistics Society*, 14, 319-339.
- Lieven, E., & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In Robinson, P., & Ellis, N. (eds.), *Handbook of*

- Cognitive Linguistics and Second Language Acquisition*, 168-196. New York, NY: Routledge.
- Mak, W., Vonk W., & Schriefers, H.. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54(4), 466-490.
- Ming, T., & Chen, L. (2010). A discourse-pragmatic study of the word order variation in Chinese relative clauses. *Journal of Pragmatics*, 42 (1), 168-189.
- Ochs, E. (1979). Planned and unplanned discourse. In Givon, T. (ed.), *Discourse and Syntax* (pp. 51-80). New York, NY: Academic Press.
- Real, F., & Christiansen, M. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, 57(1), 1-23.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford, UK: Oxford University Press.
- Sugiura, M. (2008). *Eigo gakushūsha no korokēshon chishiki ni kansuru kisoteki kenkyū [A fundamental study on English L2 learners' collocational knowledge]*. (Report of Grant-in-Aid for Scientific Research (B) (2005-2007), No.17320084).
- Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, 58 (1), 1-21.
- Thompson, S. (1997). Discourse motivations for the core-oblique distinction as a language universal. In Kamio, A. (ed.), *Directions in Functional Linguistics*, 59-82. Amsterdam, NL: John Benjamins.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Traxler, M., Morris, R., & Seely, R.. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, 47(1), 69-90.

Appendix

Examples of Relative Clause Heads and Relative Clauses

Heads

Intransitive subject (S): The man was at the store on the day.

Transitive subject (A): The girl brought her sister back home.

Object (O): She paid the bill.

Oblique (B): He cannot go to the party.

Main clause (M): He was diabetic.

Noun phrase (NP): The man.

Predicate NP (PN): There were some people there.

Relative Clauses

Intransitive subject RC (SR): ... the man who [] was in the classroom.

Transitive subject RC (AR): ... the man who [] play baseball.

Object RC (OR): ... the man I saw [].

Oblique RC (BR): ... the man I went to the school with [].