SIMULATING THE ACQUISITION OF ADVERB LICENSING: A PRELIMINARY INVESTIGATION[*]

Tomohiro Fujii
Yokohama National University

1. Introduction

The literature on English adverbs has noted that so-called Subject-Oriented adverbs, such as *wisely*, permit two readings: a clausal (or sentence-modifying) reading and a manner (or VP-modifying) reading (Jackendoff 1972; Bellert 1977; McConell-Ginet 1982; Ernst 2002).

(1) a. John wisely will leave. (Clausal reading)
    b. John will leave wisely. (Manner reading)

(1a) and (1b) can be paraphrased as (2a) and (2b), respectively.

(2) a. It is wise of John to leave.
    b. John will leave in a wise manner.

Adverbs of the relevant class receive a clausal reading in pre-auxiliary position and receive a manner reading in postverbal position. The latter restriction is lifted when a comma intonation is used. (3) requires a clausal reading since a pause is placed before the adverb.

(3) John will leave, wisely. (Clausal reading)

Additionally, Subject-Oriented adverbs generally allow both readings when they occur between the finite auxiliary and the lexical verb. (4) is two-way ambiguous.

(4) John will wisely leave. (Clausal reading/ Manner reading)

These interpretive possibilities can be captured by the generalization given in (5).

(5) Subject-Oriented adverbs are assigned a clausal interpretation when adjoined to S while they are assigned a manner interpretation when adjoined to VP.

[*] This paper grew out of a presentation given at the workshop held at Nanzan University in July 2024. I am grateful for the valuable comments provided by the audience and also benefited from discussions with Kotaro Nakamura. This work was supported by JSPS KAKENHI Grant Number JP22K00502.

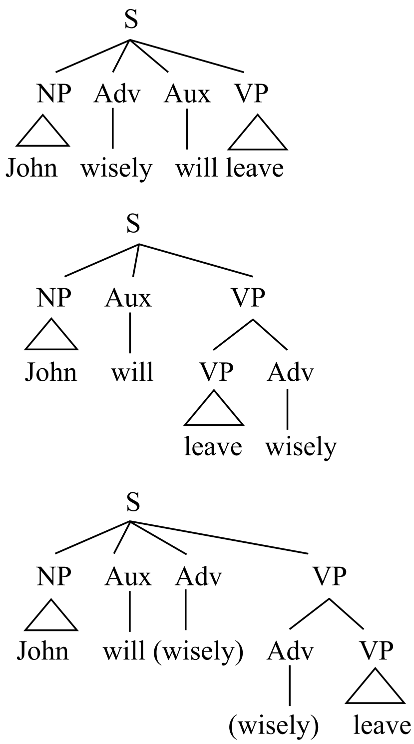
While the generalization is stated within a certain framework of sentence structure, it can be translated into a different phrase-structure assumption without difficulty; see for example Ernst (2002), where more articulated clause structure is adopted.

The diagrams in (6a-c) help to see how the generalization accounts for the form-meaning correspondence found in (1) and (4).

(6) a.

b.

c.

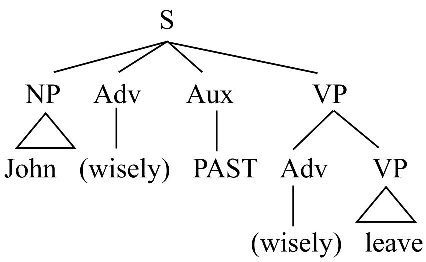


The generalization also covers the fact that sentences without an overt auxiliary such as (7) allow both clausal and manner readings.

(7) John wisely left.

After the tense affix undergoes Affix Hopping, the relative word order between the S-adjoined adverb and the verb becomes identical to that between the VP-adjoined adverb and the verb.

(8)



Against the background discussed so far, it is important to note the existence of unattested combinations of syntactic position and reading. As shown in Table 1, there are four logically possible hypotheses. Of these, two (Hypotheses 1 and 4) are attested, while the other two are not. This observation naturally raises a learnability question: how can the learner determine that the attested hypotheses are correct while rejecting the unattested ones?

Table 1. Attested and unattested combinations of positions and readings

	Interpretation	Position	Availability
Hypothesis 1	Clausal	S-adjoined	✓
Hypothesis 2	Clausal	VP-adjoined	*
Hypothesis 3	Manner	S-adjoined	*
Hypothesis 4	Manner	VP-adjoined	✓

At this point, one might wonder why examining the acquisition of adverb licensing is worthwhile. A key motivation emerges from a comparison between SVO languages like English and SOV languages like Japanese. If generalization (5) holds in, say, Japanese as well (Sawada 1978; Nakau 1980; Kubota 2015; Morzycki 2016; Ernst 2015; Miura and Fujii 2020, 2021), it follows that word order information—specifically, whether an adverb precedes the finite auxiliary of the clause or follows the lexical verb—cannot serve as a cue for the learner . This is because, in languages like Japanese, tensed verbal amalgamates appear clause-finally.

To address the broader question concerning SOV languages, the current study first tackles a more fundamental issue: can the learner select the correct generalization based on word order cues and interpretation? As will become clear below, our findings indicate that generalization (5) can be more challenging to acquire even in SVO languages like English than one might initially expect.

2. A Simulation Study

2.1. The Corpus

To investigate the distribution of *stupidly* in natural language use, we conducted a corpus search using the Corpus of Contemporary American English (COCA). Our initial search yielded 500 examples of *stupidly*, which we then filtered by removing instances that fell outside four specific uses. Excluded cases included those where *stupidly* was adjoined to a modifier (e.g., *stupidly long speech*), those where it followed an infinitive or preceded a gerund, and those where it was preceded by a comma.

After this refinement, we obtained 252 relevant examples. These were categorized into three distinct positional patterns. First, we identified 19 sentences where *stupidly* precedes a finite auxiliary, a position we refer to as *Pre-Aux* (e.g., *He stupidly had forgotten his keys*). Second, we found 96 sentences where the adverb occurs between a finite auxiliary and the lexical verb or in sentences without an overt auxiliary, which we label as *Middle* (e.g., *He had stupidly* *forgotten his keys* or *He stupidly forgot his keys*). Finally, we observed 137 sentences where *stupidly* follows a lexical verb, categorized as *Postverbal* (e.g., *He forgot his keys stupidly*). Some corpus examples are cited in (9).

(9) a. And stupidly, I thought this bank was going to let me know if I had a … (1992 SPOK ABC_Brinkley)
b. "I thought you said he was in England," I stupidly said. (1994 FIC Bk:BodyFarm)
c. ... jumping up and down, throwing rocks at you... as you crawled stupidly up onto the sand. (2001 MOV No Such Thing)

The counts observed in the corpus are summarized in Table 2.

Table 2. The frequencies of data types observed in a small corpus derived from COCA

Data type	Pre-Aux	Middle	Postverbal
Frequency	19	96	137

We can reasonably assume that Pre-Aux data points always involve clausal readings, while Postverbal data points consistently involve manner readings. We did not attempt to determine which readings the adverbs in *Middle* data points receive—a point to which we will return below.

2.2. Hypothesis Space and Its Relationship to Data Types

Table 3 illustrates how each data type aligns with the hypotheses under consideration. Hypotheses 1–4 represent all the possibilities within the hypothesis space.

Table 3. Four hypotheses and data types. The symbol "✓" indicates that the data type matches the hypothesis, while "*" indicates it does not.

	Pre-Aux	Clausal Middle	Manner Middle	Postverbal
Hypothesis 1 (Clausal, S-adjoined)	✓	✓	*	*
Hypothesis 2 (Clausal, VP-adjoined)	*	✓	*	*
Hypothesis 3 (Manner, S-adjoined)	*	*	✓	*
Hypothesis 4 (Manner, VP-adjoined)	*	*	✓	✓

The *Pre-Aux* data points, assumed to always involve clausal adverbs, are compatible only with *Hypothesis 1* (i.e., the hypothesis that the adverb is an S-adjoined clausal adverb). The *Middle* data points must be further divided into two subtypes: those with clausal readings and those with manner readings. Middle data points with clausal readings are consistent with both

Hypothesis 1 and *Hypothesis 2* (i.e., the hypothesis that the adverb is a VP-adjoined clausal adverb). In contrast, Middle data points with manner readings are consistent with *Hypothesis 3* (i.e., the hypothesis that the adverb is an S-adjoined manner adverb) and *Hypothesis 4* (i.e., the hypothesis that the adverb is a VP-adjoined manner adverb). Finally, the *Postverbal* data points, which are assumed to always carry manner readings, are compatible only with *Hypothesis 4*. Thus, the Pre-Aux and Postverbal data types are unambiguous triggers that support the correct hypotheses, while Clausal and Manner Middle data types are not.

2.3. Update Equation

The learning model we used was based on Yang's (2002, 2011) Naïve Parameter Learner. Here's how it works. First, the four hypotheses were assigned equal probabilities, each starting at 0.25. Suppose H1 (Hypothesis 1) is tested against a *Middle* data point. Since this hypothesis is consistent with the data point, its probability is increased. When a hypothesis is incompatible with a data point, it is not awarded or penalized (Yang 2011). The learning rate γ was set to be 0.05.

(10) Update equations for the Naïve Parameter Learner for a hypothesis space with four hypotheses, given a data point *d*.

a. Upon the presentation of an input datum *d*, the child selects a use H_i with the probability of P_i.

b. If H_i is compatible with *d*, $P'_i = P_i + \gamma*(1 - P_i)$

c. Otherwise, do nothing.

As their ability to update probabilities this way indicates, Naïve Parameter Learners are assumed to parse strings of words correctly and to identify the relevant aspects of the utterance meaning.

We also note that after each update, normalization ensures that the probabilities remain valid by keeping their sum equal to 1. Each probability is divided by the total sum of the four updated probabilities.

2.4. Results

Before running the simulation, it is important to note that we did not code the readings assigned to the Middle data points collected from COCA. Since no real data is available, we distributed the 96 Middle data points into two categories—Clausal Middle and Manner Middle—using two different distributions. This approach allows us to examine whether the eventual distribution of Middle data points has a significant impact on the simulation results. Simulations employed the proportions of data types found in Modified Corpus A (Table 4a), Modified Corpus B (Table 4b), and Modified Corpus C (Table 4c).

H4 (Manner, VP-adjoined) was ranked first in all simulations. This result is unsurprising, given the overwhelming frequency of the Postverbal data type. However, learning was unsuccessful when using Modified Corpus A, where Clausal Middle data points were set to be twice as frequent as Manner Middle ones. In this scenario, one of the correct hypotheses, H1 (Clausal, S-adjoined), failed to survive the tests. This outcome appears to be due to the Clausal Middle data type, the second most frequent category, partially supporting H2 (Clausal, VP-adjoined). As a result, H2 gained strength at the expense of the remaining two hypotheses, H1 (Clausal, S-adjoined) and H3 (Manner, S-adjoined). Figure 1 shows the results of simulations.

Table 4. The frequencies of data types with a modification.

(a) Modified Corpus A. *Clausal Middle* data points are assumed to occur twice as frequently as Manner *Middle* data points.

Data type	Pre-Aux	Clausal Middle	Manner Middle	Postverbal
Frequency	19	64	32	137
Percentage	8 %	25 %	13 %	54 %

(b) Modified Corpus B. *Manner Middle* data points are assumed to occur twice as frequently as *Clausal Middle* data points.

Data type	Pre-Aux	Clausal Middle	Manner Middle	Postverbal
Frequency	19	32	64	137
Percentage	8 %	13 %	25 %	54 %

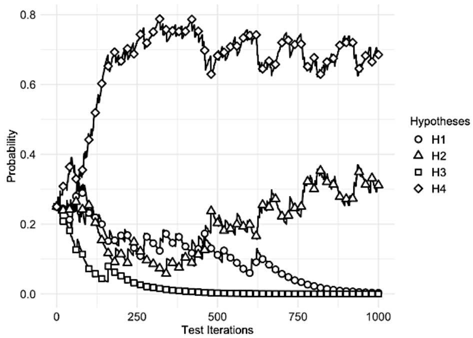
(c) Modified Corpus C. *Manner Middle* data points are assumed to occur as frequently as *Clausal Middle* data points.

Data type	Pre-Aux	Clausal Middle	Manner Middle	Postverbal
Frequency	19	48	48	137
Percentage	8 %	19 %	19 %	54 %

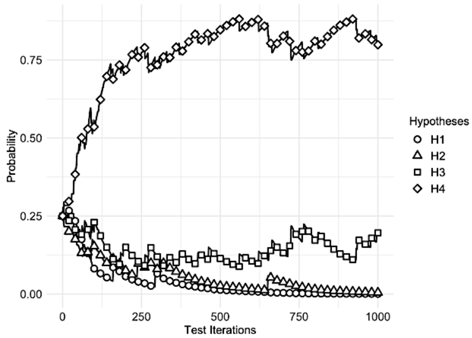
A virtually parallel situation occurred in the simulation based on Modified Corpus B. In this case, the second most frequent category was the Manner Middle data type, which partially supported H3 (Manner, S-adjoined). Consequently, this weakened H1 and H2.

Finally, only the simulation based on Modified Corpus C yielded a successful outcome by eliminating the incorrect hypotheses, H3 and H4. What explains this success? We believe there are two key reasons: (i) the two Middle data types were equally frequent, preventing H2 and H3 from outcompeting one another; and (ii) the Pre-Aux data type, which unambiguously supported H1, was frequent enough to prevent it from being overshadowed by H2 or H3.

Modified Corpus A



Modified Corpus B



Modified Corpus C

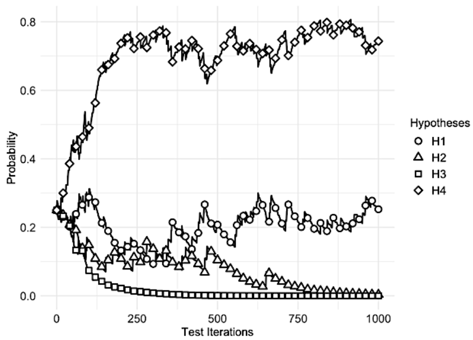


Figure 1. Probability updates with 1000 data points based on Modified Corpora A, B and C. In Corpus A, Clausal Middle data points were assumed to be twice as frequent as Manner Middle data points. In Corpus B, the opposite was assumed to hold. In Corpus C, the two data types were assumed to be equally frequent.

3. General Discussion and Conclusion

This study preliminarily proposed a learning model for the form-meaning correspondence observed with Subject-Oriented adverbs in English, using a Naïve Parameter Learner for usage selection. Although the analysis was based on a small corpus focusing on a single adverb, the following general picture emerges: H1 is incorrectly eliminated if (i) the frequency of the Postverbal data type exceeds that of the other data types, (ii) the frequency of the Middle data type surpasses that of the Pre-Aux data type to a certain extent, and (iii) the distribution of clausal and manner interpretations within the Middle data points is unbalanced.

If this happens to the input often enough, then the model is unlikely to approximate the actual learning model that humans internalize. What aspect of the model caused this issue? The hypothesis space underlying the model has one fundamental property: meaning and form are paired arbitrarily, without adhering to the Fregean/Montagovian principle of compositionality (Goldberg 1995). This lack of compositionality—namely, the pairing of the adverb's two readings with the two nodes to which it is adjoined—is what allows the hypothesis space to include H2 and H3. Although we cannot determine what a learning process incorporating compositionality would look like, or whether such a model would successfully select the correct hypotheses, the current results provide support for the relevance of compositionality.

References

Bellert, Irena (1977) "On semantic and distributional properties of sentential adverbs." *Linguistic Inquiry* 8, 337-351. https://www.jstor.org/stable/4177988.

Davies, Mark (2008-) *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Ernst, Thomas (2002). *The Syntax of Adjuncts*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486258.

Ernst, Thomas (2015) "Evidence for a proper treatment of the clausal/manner distinction: Comments on Kubota, Transforming manner adverbs into surface-subject-oriented adverbs: Evidence from Japanese." *Natural Language & Linguistic Theory* 33: 1047–1055. https://www.jstor.org/stable/43697795.

Goldberg, Adele (1995) *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: Chicago University Press.

Jackendoff, Ray (1972) *Semantic Interpretation in Generative Grammar*. Cambridge, Mass: MIT Press.

Kubota, Ai (2015) "Transforming manner adverbs into subject-oriented adverbs: evidence from Japanese." *Natural Language & Linguistic Theory* 33: 1019–1046. http://www.jstor.org/stable/43697794.

McConnell-Ginet, Sally (1982) "Adverbs and logical form: a linguistically realistic theory." *Language* 58: 144-184. https://www.jstor.org/stable/i217178.

Miura, Kaori and Tomohiro Fujii (2020) "Highs and lows of the syntax of agent-oriented adverbs in Japanese." In Tae Sik Kim and Sae-Youn Cho (eds.), *Proceedings of the 22nd Seoul International Conference on Generative Grammar (SICOGG)*. 188–200.

Miura, Kaori and Tomohiro Fujii (2021) "Japanese subject-oriented adverbs in a scope-based theory of adverbs." *Proceedings of the Linguistic Society of America* 6 (1), 254–264.

Morzycki, Marcin (2016) *Modification*. Cambridge, UK: Cambridge University Press.

Nakau, Minoru (1980) "Bunhukusi no hikaku (A contrastive study of sentence adverbials)." In Tetsuya Kunihiro, ed. *Nitieigo Hikaku Kooza 2, Bunpoo [A comparative study of Japanese and English 2: Grammar]*, 157–219. Tokyo: Taishukan.

Sawada, Harumi (1978) "Nichieigo bunfukusirui no taishoo gengogakute kikenkyuu [A contrastive study of sentence adverbials in English and Japanese]." *Gengo Kenkyu* 74: 1–36.

Yang, Charles (2002) *Knowledge and Learning in Natural Language*. Oxford University Press.

Yang, Charles (2011) "Computational models of syntactic acquisition." *WIREs Cognitive Science*. doi: 10.1002/wcs.1154