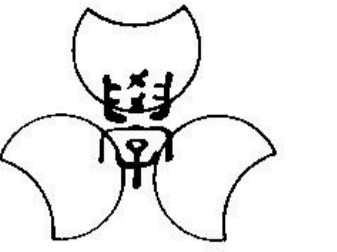


# Why 'What do you like sports?': Evidence for Shallow Structures



Ralph L. Rose, Waseda University School of Science and Engineering  
Natsuki Amano, Gunma Prefectural Women's University



## Abstract

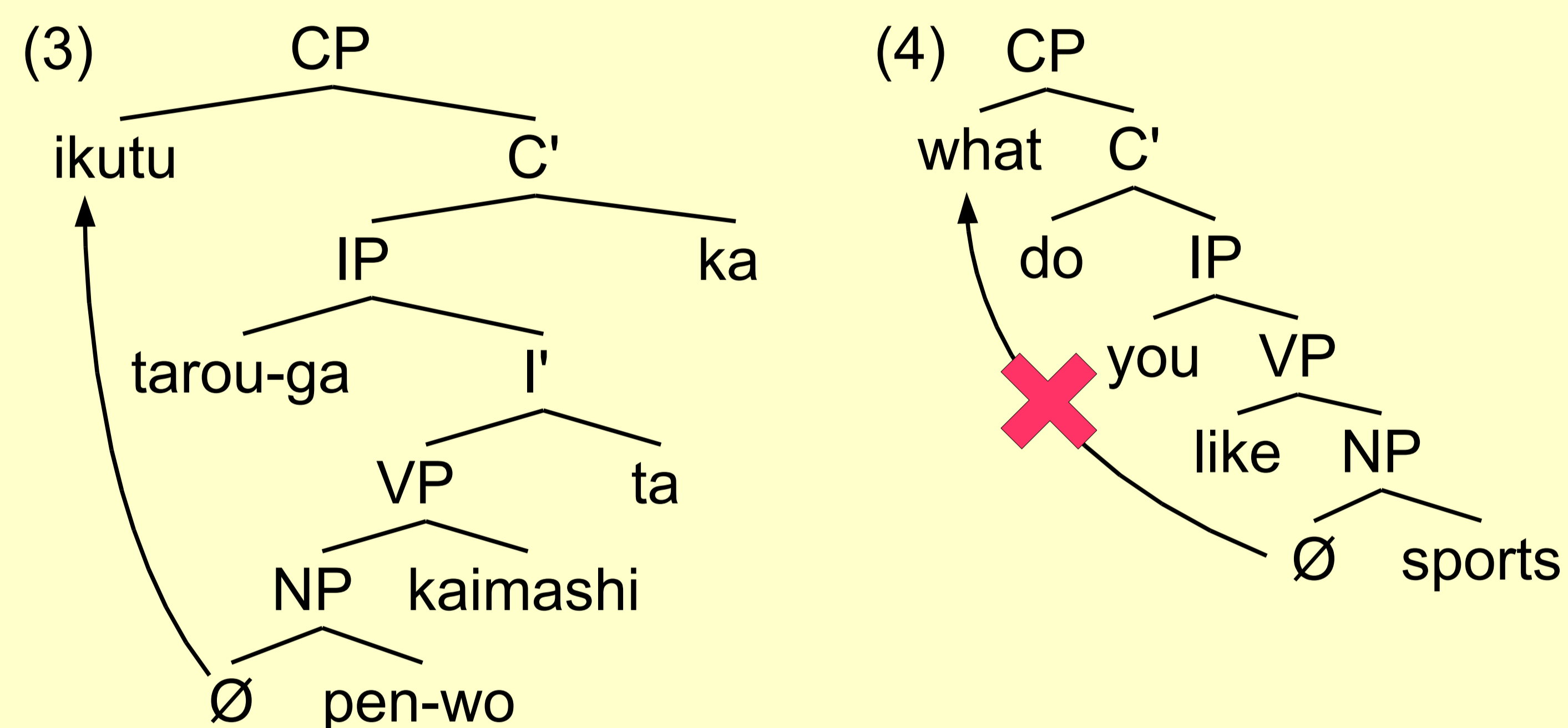
One explanation for the commonly observed mistake by early Japanese learners of English—\**What do you like sports?*—assumes that the error is a transfer error: Japanese allows left-branch extraction while English does not (cf., Yamane, 2003). A potentially simpler explanation is that learners are creating shallow structures based on memorized chunks of language: [*what do you like*] [*sports*]. In order to test the different predictions of these two explanations, two pilot studies were done: a corpus study and a psycholinguistic experiment. Results of both studies are more consistent with a shallow structures explanation.

## Background

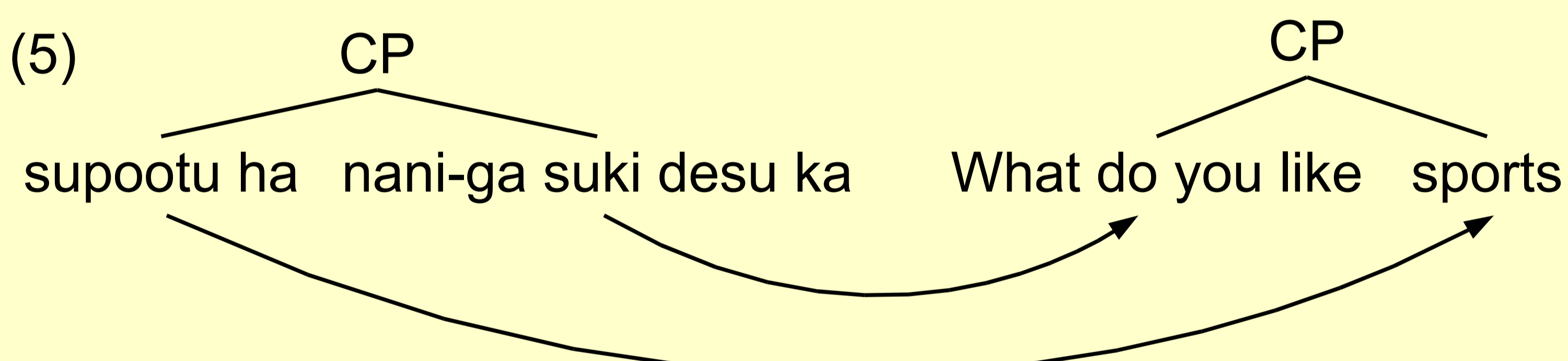
One common error in question formation by Japanese speakers of English is shown in (1) where the target grammatical form is (2).

- (1) \*What do you like sports?
- (2) What sports do you like?

This can be explained as the result of transfer (cf., Yamane, 2003). In Japanese, extraction of a wh-word out of the left branch of a constituent is allowed as shown in (3). In English, however, such left-branch extraction is not allowed as shown in (4). In this explanation, learners transfer the UG parameter setting for left-branch extraction from Japanese into English, thereby permitting the production of (1).



One question is whether second-language learners in general actually compute such complex structures or merely produce shallow structures (cf., Clahsen and Felser, 2006) based on memorized chunks. Based on this, an alternative and simpler account (hereafter shallow structures account) would be that the learner's representation of (1) could be a simple two-branch structure derived as shown in (5).



If this account is correct, then the synonymous question in (6) should lead to fewer errors because the shallow chunk-based analysis more closely matches English word order.

- (6) nan-no spootu ga suki desu ka

Previous experimentation has not controlled whether participants underlyingly think of the structure in (5) or (6). Thus, there is a confound relating to whether *sports* can be topicalized as in (5) or not topicalized as in (6). The present study seeks to control this in order to evaluate the complex structures versus the shallow structures accounts.

## References

- Clahsen, H. and Felser, C. (2006). Grammatical processing in language learning. *Applied Psycholinguistics*, 27:3-42.
- Kellerman, E. (1985). *Input in Second Language Acquisition*. Rowley, MA: Newbury House.
- Tono, Y. (2007). 日本人中高生一万人の英語コーパス [English Corpus of 10,000 Japanese Junior and Senior High School Students]. Tokyo, Japan: Shogakkan.
- Yamane, M. (2003). On Interaction of First-Language Transfer and Universal Grammar in Adult Second Language Acquisition: Wh-Movement in L1-Japanese/L2-English Interlanguage. Ph.D. Dissertation. University of Connecticut.

## Corpus Study

Sentences with a (intended) *what*-constituent were extracted from the Japanese EFL Learner Corpus (Tono, 2007).

	Preferred Center (wh-complement may not topicalize)	Not Preferred Center (wh-complement may topicalize)
Control (well-formed)	[description of writer's dreams] ... I forgot <b>what dream</b> I had on the last day of 1999 .	[description of rice] ... I like considering <b>what sort fo dishes</b> match rice.
	7	1
Split (not well-formed)	[description of writer's dreams] ... <b>What</b> did you see bad <b>dream</b> ?	[description of family meals] ... <b>what</b> do you have <b>きれいなもの</b> ?
	2	5

$\chi^2(1) = 5.4, p < 0.05$   
(note: because of low counts, this is not reliable)

Results are not predicted by the complex structures account but are consistent with the shallow structures account.

## Psycholinguistic Experiment

### Design/Materials

The experiment employs a design in which participants were asked to imagine themselves as interpreters during an interview. Their task was to translate questions into English for a pre-recorded interviewee. Thus, dialogs were prepared as follows using non-topicalized and topicalized Japanese questions (in red).

Not Topicalized		Topicalized	
Interviewer:	スポーツは好きですか。	Interviewer:	週末は何をしますか。
Interpreter <sub>1</sub> :	Do you like sports?	Interpreter <sub>1</sub> :	What do you do on weekends?
Interviewee:	Yes, I do.	Interviewee:	I read books.
Interpreter <sub>1</sub> :	はい、そうです。	Interpreter <sub>1</sub> :	本を読みます。
Interviewer:	<b>何のスポーツが好きですか。</b>	Interviewer:	<b>スポーツは何が好きですか。</b>
Interpreter <sub>2</sub> :	?	Interpreter <sub>2</sub> :	?
Interviewee:	I like tennis.	Interviewee:	I like tennis.

A total of 18 dialogs were prepared: 6 with *nani*, 6 with *ikura/ikutu*, and 6 with *doko* (which/where). Participants listened to the dialogs and played the role of Interpreter<sub>2</sub>, providing spoken interpretations.

### Results

Participants included 53 native speakers of Japanese studying English at Gunma Prefectural Women's University. The participants were divided into two groups (HIGH, LOW) based on TOEIC scores ( $\mu_{HIGH} \approx 700, \mu_{LOW} \approx 510$ ) because lower-level learners are believed to rely more on memorized chunks (cf., Kellerman, 1985). Responses were scored as showing left-branch extraction or not.

GROUP	CONTEXT		CONTEXT	By subjects n.s.	By items n.s.
	Not Topicalized	Topicalized			
High	3.4%	3.1%	F(1,51)=2.9, p=0.09		
Low	5.2%	8.6%	F(1,51)=2.6, p=0.11	F(1,17)=5.0, p<0.05	F(1,17)=5.7, p<0.05

Results also vary across different wh-words.

- ✓ *nani*: main effect of CONTEXT
- ✓ *ikura/ikutu*: no effects
- ✓ *doko*: main effect of GROUP and CONTEXT; marginal interaction

## Discussion/Conclusion

The results are surprising because they are not predicted by the complex structures account discussed above. Rather, they are more consistent with the alternative shallow structures (i.e., chunk-based translation) account. However, the effect is more clearly seen with the LOW learners. But in fact, these learners are relatively high: The effect may be better observed with much lower level learners (say, high school students).

The differences among the three wh-words is also at odds with the complex structures account which would predict no difference among all of these items. The differences might be explained by the shallow structures model and by taking into account the influence of how and when the various chunks have been learned by students.

In summary, The experimental results point toward a shallow-structures model of second language production, but further work is necessary to confirm these conclusions.