Quantifier scope in non-native Japanese : a comparative interlanguage study of Chinese-, English- and Korean-speaking learners

Marsden, Heather Lynn

Quantifier scope in non-native Japanese: A comparative interlanguage study of Chinese, English, and Korean-speaking learners

Heather Lynn Marsden

PhD dissertation

School of Linguistics and Language
University of Durham

2004

Marsden, Heather. 2004. *Quantifier scope in non-native Japanese: A comparative interlanguage study of Chinese, English, and Korean-speaking learners*. PhD dissertation, University of Durham.

Errata

Section 3.2.4, pp. 111-117

Dekydtspotter, Sprouse, Swanson & Thyre (1999) is a pilot study for Dekydtspotter, Sprouse & Swanson (2001), therefore the two papers do not report on a single dataset, contrary to the claim in this section. However, the two papers *do* both investigate the phenomenon described in this section, and the results for DS&S (2001) are very similar to those for DSS&T (1999).

Footnote 149, p. 186

Every instance of *enu N-to* should be *enu N-(i)na*. The corrected footnote is as follows:

Korean also has the universal QP *enu N-(i)na* 'every N', which is morphologically, syntactically and semantically similar to Japanese *dono N-mo*. *Enu N-(i)na* was not included in the Korean version of the experimentation due to insufficient knowledge of this QP at the time of having the opportunity to work with native Korean test participants. To gain at least pre-experimental data on *enu N-(i)na*, Korean versions of Types 1a–b and 3a–b were created (with *enu N-(i)na* as the object QP) and presented to two native Korean linguists. They reported that object-wide scope was unacceptable, as expected following the theoretical literature. (Beck & Kim (1997: 374) discuss one example with *enu N-(i)na* 'every N' as object QP; other sources, such as Kim 1989, discuss only *nwukwuna* 'everyone'.)

Abstract

This thesis investigates native language (L1) influence and innate linguistic knowledge (i.e., Universal Grammar) in non-native language (L2) acquisition by means of a comparative interlanguage study of quantifier scope interpretation in L2 Japanese, by adult native speakers of English, Chinese, and Korean. The phenomena investigated are:

i. the availability of object-wide scope in sentences with an existentially-quantified subject and universally-quantified object (e.g., *Someone read every book.*);

ii. the availability of a pair-list reading in questions with *everyone* as the subject and *what* as the object (e.g., *What did everyone buy?*).

Picture-sentence match tasks are developed to investigate these two phenomena in native Japanese, English, Chinese, and Korean, as well as in English/Japanese Chinese/Japanese and Korean/Japanese interlanguage. The native experimental data confirm that, with respect to (i), the object-wide scope interpretation ('for each book, someone read it') is readily available in English but not in Japanese, Chinese, or Korean; and with respect to (ii), a pair-list answer (e.g., *Sam bought apples, Jane bought pears, Sue bought...*) is readily available in English, Chinese and Korean, but not in Japanese.

These cross-linguistic differences are exploited in the investigation of two main predictions based on Schwartz & Sprouse's (1994, 1996) Full Transfer/Full Access model of L2 acquisition: (1) the L2 learner groups will show divergent development with respect to Japanese scope interpretation due to the distinct scope interpretation possibilities in their respective L1s; (2) advanced L2 learners of Japanese will demonstrate native-like knowledge of quantification phenomena even under severe poverty of the stimulus, due to L2 acquisition being constrained by UG. The results support both predictions. On the basis of these findings, it is concluded that both the L1 and UG are privileged sources of knowledge in the L2 acquisition of phenomena at the syntax-semantics interpretive interface.

Acknowledgements

I would like to thank everyone who helped to make this dissertation possible.

First, thanks are due to my supervisors, Anders Holmberg and Bonnie Schwartz, for careful reading and discussion of drafts, and for much-appreciated encouragement.

Second, I would like to thank all those who contributed towards completion of the experimental work. I am grateful to Sae Kawase, Ritsuko Koso, Makiko Mukai, and Mari Umeda for help with proofreading and checking the Japanese test materials and recording the audio components; to Yannan Guo, Limin Jin, Yuet-Wah Lam, and Nianling Yang for help with translating the test materials into Chinese and recording the Chinese audio components; to Jin Hee Park for help with translating the test materials into Korean and recording the Korean audio components; to Joseph Emonds, Makiko Hirakawa, Jane Ives, Sandra McNulty, Kayoko Matsuo, Shigeo Miyamoto, Kizuki Morimoto, Barbara Pizziconi, Ayako Sakata, Eleanor Spivey, Isoo Tabuchi, David West, Melinda Whong-Barr, He Ying, and Hitoko Yoshikawa for recruiting test participants and providing venues for conducting the tests; and to all the individuals who gave their time to participate in the tests themselves. In addition, I am grateful to Toshiko Ishida for discussion of Japanese language testing, and to Simon Liversedge and Dan Robertson for invaluable advice on statistical analysis.

Further thanks are due to Seiki Ayano, Naoko Hamada, Mayumi Hosono, Sae Kawase, Mika Kizu, Ritsuko Koso, Makiko Mukai, Hitomi Nakata, Yoko Nishina, Kanako Nishizumi, Kyoko Oga, Kyoko Otsuki, Hideki Saigo, Tomoko Semba, Hidekazu Tanaka, Mari Umeda, and Kyoko Verran for patient discussion of judgements about quantifier scope interpretation in Japanese; to Nonie Chang, Yu Jiang, Yuet Wah Lam, Li-Chen Li and Zhengzheng Wang for the same in Chinese; and to Jin Hee Park and Kook-Hee Gill for the same in Korean.

I am grateful to the following linguists for helpful discussion of issues of quantification, and for sending me papers: Laurent Dekydtspotter, Nobuko Hasegawa, Ringe Hayashishita, Kook-Hee Gill, Susumu Kuno, Thomas Hun-tak Lee, Ayumi Matsuo, Yoichi Miyamoto, Tetsuya Sano, Rex Sprouse, Tim Stowell, Satoshi Tomioka, Noriaki Yusa.

I acknowledge with thanks the support provided by an award from the Japan Foundation Endowment Committee (502214) for data collection conducted in Japan.

Finally, I thank Lewis for invaluable help with the meaning of everything.

Table of Contents

1. Introduction

Two key questions animating research into adult non-native language (L2) acquisition during the past twenty years are as follows:[1]

1. Is (adult) L2 acquisition constrained by the same innate linguistic mechanisms (i.e., Universal Grammar, UG) as first language (L1) acquisition?

2. What is the role of L1 knowledge in (adult) L2 acquisition?

The goal of this dissertation is to present quantitative, empirical data that bear on these questions, by investigating an area that has, as yet, received little attention in L2 acquisition research: namely, L2 knowledge of phenomena at the syntax-semantics interface. Specifically, the study investigates quantifier scope interpretation in L2 Japanese by learners whose L1s are English, Korean, or Chinese.

Schwartz & Sprouse (2000) point out that the conclusions of existing research addressing the questions in (1) and (2) are often jeopardised by changes in the linguistic theory within which investigations are set. Since the precise architecture of UG is not yet known, L2 acquisition research that is intimately tied to the technicalities of a particular linguistic hypothesis risks becoming invalid when that hypothesis is revised. Schwartz & Sprouse (2000) advocate the following methodological framework as a means of addressing (1) and (2) while bypassing the problem of changing linguistic theory: (i) investigation of clear L2 poverty-of-the-stimulus problems, in order to identify the role of UG in L2 acquisition; and (ii) comparison of the developmental paths of learners whose L1s are typologically distinct with respect to a specific target language phenomenon, in order to identify the role of L1 transfer. This framework is adopted for the present study.

The advantages of the framework, as detailed by Schwartz & Sprouse (2000), are as follows. First, investigation of L2 poverty-of-the-stimulus phenomena is the only sure-fire way to discover whether L2 acquisition is constrained by UG. L2 poverty of the stimulus is defined as underdetermination of a target language phenomenon by the target language input, by L1 knowledge, and (where L2 learners

[1] The term 'L2 acquisition' is used throughout this dissertation to cover all non-native language acquisition, whether or not the 'L2' is strictly a 'second' language or a third (or fourth, fifth, etc.). Henceforth, 'L2 acquisition' refers to L2 acquisition by adult learners. 'Adult', in this context, refers to post-puberty.

['L2ers'] are instructed) by classroom teaching. The utility of such phenomena in L2 acquisition research was identified in the 1980s by Lydia White (e.g., White 1989a, 1989b). As White (1989b: 46) states, '[i]f learners attain knowledge which could not have come via their mother tongue, and which could not have been induced from the input alone, arguments for a role for UG in L2 acquisition are strengthened'. Thus, evidence of L2 knowledge that goes beyond the input and L1-derived knowledge (and any relevant meta-linguistic knowledge via instruction) implies that L2 acquisition is constrained by UG, since, logically there is no other source for such knowledge. Such evidence could not be undermined by a change in linguistic theory, because a poverty-of-the-stimulus problem is a poverty-of-the-stimulus problem whatever the linguistic theory that accounts for the relevant phenomena.

Regarding the second part of the framework, Schwartz & Sprouse (2000: 181) spell out the logic of comparative interlanguage research as follows:[2]

3. In the acquisition of some phenomenon *P* in a given Target Language, compare the developmental paths of L2ers whose L1s are, with respect to *P*, typologically distinct. If one finds divergence in developmental paths [...], one has evidence for transfer in that domain—because there is nothing in the L2ers' input [...] which could account for such divergence. If, on the other hand, one finds a uniform developmental path with respect to *P*, one has evidence against transfer.

As with investigation into L2 poverty-of-the-stimulus problems, the results of such a study are independent of the linguistic theory accounting for the cross-linguistic differences.

The general research questions given in (1) and (2) are restated in (4) and (5) in terms of Schwartz & Sprouse's (2000) methodological framework.

4. When a target language phenomenon *P* represents an L2 poverty-of-the-stimulus problem, are L2ers able to overcome the problem and acquire *P*?

5. Do L2ers show divergence with respect to *P* when their L1s are typologically distinct with respect to *P*?

[2] The term 'interlanguage' refers to the entirety of the L2 learner's internalized knowledge of the target language. This knowledge can be thought of as a continuum between the initial-state L2 knowledge and the final state. At any point along the continuum the knowledge is assumed to be systematic and rule-governed. (See Larsen-Freeman & Long 1991: 60; White 1989: 35-36.)

If L2 data provide an affirmative answer to (4), this implicates UG in the L2 acquisition process. If they provide an affirmative answer to (5), this implicates L1 knowledge as a determinant of L2 development.

The questions in (4) and (5) are the research questions motivating the present study. To address (4), two L2 poverty-of-the-stimulus phenomena are identified in the L2 acquisition of Japanese quantifier scope interpretation. Briefly, these are (i) native English speakers' acquisition of the unambiguity of Japanese doubly-quantified sentences ('QP-QP sentences')[3] such as (6a): specifically, that these allow only a subject-wide scope interpretation (6b) and not an object-wide scope interpretation (6c) (Hoji 1985; Kuroda 1970); and (ii) native English and Chinese speakers' acquisition of the unambiguity of scrambled Japanese [*wh*-object...QP-subject...] questions, such as (7a): specifically, that they allow only an 'individual answer' (7b), and not a 'pair-list answer' (7c) (Hoji 1985).[4]

6. a. Dareka-ga dono hon-mo yonda.
   someone-NOM every book read
   'Someone read every book.'

   Interpretation of (6a):
   b. There is some person x such that x read every book. (Subject-wide scope)
   c. *For every book y, there is some person who read y. (Object-wide scope)

7. a. Nani-o daremo-ga katta no.
   what-ACC everyone-NOM bought Q
   'What did everyone buy?'

   Example answers to (7a):
   b. 'A book.' (Individual answer)
   c. *'Bill bought a book, Sally bought a pen, Jane bought a bag, ...' (Pair-list answer)

Acquisition of the lack of object-wide scope interpretation of (6a) represents a poverty-of-the-stimulus problem for learners whose L1 is English because the absence of this reading is underdetermined by the following:

8. a. the L1: equivalent English sentences allow an object-wide scope interpretation;

[3] The abbreviation 'QP' ('Quantified Phrase') is used to refer to quantified NPs (e.g., *everyone, every woman*) and PPs (e.g., *in every shop*).
[4] These quantifier scope interpretation phenomena, along with the L2 acquisition problems they entail, are described in more detail in Chapter 2.

b. the target language input: failure to encounter sentences like (6a) in an object-wide scope context does not constitute evidence that Japanese QP-QP sentences are never used in such a context; and
c. classroom teaching: commonly used Japanese language textbooks do not cover the facts of (6).[5]

Similarly, acquisition of the lack of pair-list answers in scrambled Japanese [*wh*-object...QP-subject...] questions such as (7a) is underdetermined for L1 English- and L1 Chinese-speaking learners of Japanese by the following:

9. a. the L1s: English and Chinese allow pair-list interpretations (see, e.g., Aoun & Li 1993);
b. the target language input: lack of pair-list answers in the input cannot motivate induction of the fact that Japanese grammar does not allow pair-list interpretations; and
c. classroom instruction.[6]

An important body of research into the L2 acquisition of subtle syntax-semantics interface phenomena under poverty of the stimulus has already been conducted in English/French interlanguage by Laurent Dekydtspotter and colleagues (e.g. Dekydtspotter, Sprouse & Swanson 2001; Dekydtspotter, Sprouse, Swanson & Thyre 1999; Dekydtspotter, Sprouse & Thyre 1998; see Chapter 3). Findings from these studies indicate that native-like knowledge can arise, despite profound poverty of the stimulus. The present research expands Dekydtspotter *et al.*'s database on L2 acquisition at the interpretive interface by investigating a different L2, namely Japanese, and, moreover, by investigating the role of L1 influence.

To address the question of L1 influence, as articulated in (5), two comparisons were planned: learners whose L1s are Chinese or Korean with learners whose L1 is English, with respect to the interpretation of Japanese QP-QP sentences (such as (6a)); and learners whose L1 is Korean with learners whose L1 is English or Chinese, with respect to the interpretation of scrambled Japanese [*wh*-object...QP-subject...] questions (such as (7a)). In the first comparison, Chinese and Korean are distinct from English in QP-QP interpretation, in that Chinese and Korean, like Japanese, exhibit scope rigidity (i.e., they do not allow the inverse scope interpretation in [S…O…] QP-QP sentences; see Aoun & Li 1993; Kim 1989), in

[5] Textbooks consulted include the *Japanese for Busy People* series (AJLT 1996/1997), the *Yookoso!* series (Tohsaku 1994, 1995) and *An introduction to modern Japanese* (Bowring & Laurie 1992).
[6] See footnote 5.

contrast to the scope ambiguity exhibited in English. The second comparison was based on claims that Korean, like Japanese does not allow pair-list interpretations in questions with a *wh*-object and QP-subject, in contrast to Chinese and English (Beck & Kim 1997; Kim 1989).[7]

The L2 poverty-of-the-stimulus problems and interlanguage comparisons thus described are investigated in the context of Schwartz & Sprouse's (1994, 1996) Full Transfer/Full Access model of L2 acquisition. This model hypothesises that the initial state of L2 acquisition is characterised by transfer of the entirety of the L1 grammar to the interlanguage. Restructuring of this L1-based interlanguage grammar is motivated by the failure of the current interlanguage state to represent the target language input, and, moreover, the successive restructurings are hypothesised to be fully constrained by UG. The full availability of UG means that L2 poverty-of-the-stimulus problems may be overcome, since whatever UG mechanisms guide the construction of the L1 grammar with respect to a given phenomenon are also accessible in L2 acquisition. If the right (indirect) trigger is present in the target language input—and it is not obscured by the L1 grammar—the L1-based interlanguage grammar could be restructured.[8]

Experimental hypotheses are set up to test Full Transfer/Full Access as follows:[9]

10. *QP-QP interpretation:*
    a. *L1 Transfer*:
    Due to L1 transfer, lower proficiency learners of Japanese whose L1 is English will accept non-native-like object-wide scope in Japanese SOV QP-QP sentences; however, lower proficiency learners of Japanese whose L1 is Chinese or Korean will reject non-native-like object-wide scope.

    b. *UG Access*:
    Due to access to UG, advanced learners of Japanese whose L1 is English will reject non-native-like object-wide scope in Japanese SOV QP-QP sentences.

---

[7] This comparison is modified in Chapter 5, due to robust quantitative data from native-Korean speakers showing that pair-list answers are acceptable in Korean, contra the claims of the literature.
[8] Thus, the crucial difference between L1 acquisition and L2 acquisition, under Full Transfer/Full Access, is the difference in initial states: in L1 acquisition all UG constraints are either unset or set to a default at the outset; in L2 acquisition, all UG constraints are set as for the L1 at the outset. Consequently, as Schwartz (1998: 148) clarifies, '[it is] precisely because (i) UG and learnability principles [...] are constant across L1 and L2 acquisition of [language] *L* but (ii) their **initial states** are distinct, the 'final states' of L2 acquisition of *L* do not systematically replicate the final state of L1 acquisition of *L*.'
[9] These hypotheses are refined in Chapter 5.

11. *Wh-QP interpretation:*
    a. *L1 Transfer*:
    Due to L1 transfer, lower proficiency learners of Japanese whose L1 is English or Chinese will accept non-native-like pair-list answers to scrambled Japanese [*wh*-object...QP-subject...] questions: however lower proficiency learners of Japanese whose L1 is Korean will reject non-native-like pair-list answers.

    b. *UG Access*:
    Due to access to UG, advanced learners of Japanese whose L1 is English or Chinese will reject non-native-like pair-list answers to scrambled Japanese [*wh*-object...QP-subject...] questions.

The hypotheses are tested by means of a picture-matching acceptability judgement task, which is developed through careful pilot-testing. L2 Japanese data are collected from learners whose L1s are English, Chinese, or Korean. In addition, control data are collected in native Japanese, native English, native Chinese, and native Korean.

2. Quantifier scope interpretation in Japanese, Chinese, English and Korean: theories, and implications for L2 acquisition

2.0. Introduction

This chapter has two main aims: (i) to present an overview of theoretical approaches to quantifier scope phenomena in Japanese, Chinese, English and Korean;[10] and (ii) to explore the implications of each approach for the L2 learnability of Japanese quantifier scope interpretation by native speakers of Chinese, English and Korean. Section 1 comprises an outline of the key cross-linguistic differences in quantifier scope interpretation between the four languages under investigation, followed by a summary of the L2 acquisition problems posed by the Japanese quantification facts for learners whose L1s are Chinese, English, or Korean. The subsequent sections present an overview of a number of important theories of quantifier scope, focusing particularly on how they address the cross-linguistic differences presented in Section 1. The theories are divided into two categories: Section 2 presents those that treat all quantifiers in a uniform way, while Section 3 presents those that attempt to account for the different scopal properties of different quantifiers (e.g., *every* v. *all*). Section 2 thus introduces the theories of May (1977, 1985), Hoji (1985), Aoun & Li (1989, 1993) and Hornstein (1995). Section 3 presents the account by Beghelli (1995, 1997) and Beghelli & Stowell (1997), then introduces a number of accounts—including Kuno & Takami (2002), and Hayashishita (1999, 2000a, 2000b)—that focus on the roles of lexical semantics, pragmatics, and other factors in scope interpretation, in contrast to the syntax focus of the previous accounts. The accounts of May (1977, 1985), Hoji (1985) and Aoun & Li (1989, 1993) are set within the *Government and Binding* (Chomsky 1981) and *Barriers* (Chomsky 1986) frameworks of syntactic theory, while Hornstein (1995) and Beghelli (1995) are Minimalist (Chomsky 1993) approaches to quantifier scope. Kuno & Takami (2002) present a functional account, while Hayashishita (1999, 2000a, 2000b) offers a combined syntactic and cognitive analysis of scope interpretation. Each account is outlined, briefly evaluated, and discussed with respect to how learners of Japanese could acquire the characteristics

[10] The term 'quantifier' refers throughout to quantificational determiners (e.g., *every*, *some*, *a*, numerals). Quantificational adverbs (e.g., *always*, *usually*) are not discussed.

of Japanese scope interpretation under that account. Finally, Section 4 summarises the chapter.

2.1. Quantifier scope in Japanese, English, Korean, and Chinese

Hoji (1986: 95) notes two key characteristics of quantifier scope interpretation in doubly-quantified sentences ('QP-QP sentences') in Japanese that differentiate it from English:

12. Scope interpretation is more restricted in Japanese than in English in the sense that it reflects more closely the hierarchical relation[s] at S[urface]-structure.

13. In Japanese the *wh*-phrase always takes wide scope over a non-*wh* quantifier; but English does not have this restriction.

These two characteristics form the focus of investigation in this dissertation. The first is exemplified by comparing the interpretations available in (14) and (15):

14. Dareka-ga daremo-o semeta.
    someone-NOM everyone-ACC criticised
    'Someone criticised everyone.' (Hoji 1995: 336 (55a))
    Interpretation: There is some person who criticised everyone.
    $\exists x \forall y$ criticise(x,y)

15. Someone criticised everyone.
    Interpretation: i. There is some person who criticised everyone.
    $\exists x \forall y$ criticise(x,y)
    ii. For every person, there is some person who criticised him/her. $\forall y \exists x$ criticise(x,y)

The Japanese QP-QP sentence in (14) allows only a linear scope interpretation, whereby the subject quantifier takes scope over the object quantifier. The English equivalent in (15) additionally allows an *inverse* scope interpretation (15ii), whereby the object quantifier takes inverse scope over the subject quantifier. Thus scope interpretation in the canonical Japanese SOV sentence is more restricted than in SVO English.

However, inverse scope in Japanese is not ruled out completely. As observed by Kuroda (1970) and Kuno (1973), when a quantified object is scrambled over a quantified subject in Japanese, the sentence becomes ambiguous. Thus the scrambled equivalent of (14), given in (16), has both the subject-wide (S>O) and the object-wide (O>S) interpretations, as indicated:[11, 12]

16. Daremo$_i$-o dareka-ga t_i semeta *Interpretation*: S>O; O>S
everyone-ACC someone-NOM criticised
'Someone criticised everyone. (*scrambled*)' (Hoji 1995: 342 (68b))

In short, the following generalisation obtains for Japanese (Hoji 1985: 241):

17. a. QP-NOM QP-ACC V. unambiguous: S>O; *O>S
b. QP$_i$-ACC QP-NOM t_i V. ambiguous: S>O; O>S

Hoji's second characteristic of quantifier scope in Japanese—that a Japanese *wh*-phrase always takes wide scope over non-*wh* quantifiers (13)—is illustrated by means of the paradigm in (18) (Hoji 1985: 269–270, (122a–d)):[13]

18. a. *SOV [Wh-subject... QP-object...] question:*
Dare-ga daremo-o syootaisimasita ka.
who-NOM everyone-ACC invited Q
'Who invited everyone?'
Interpretation: *wh*>QP, 'For which person did that person invite everyone?'
*QP>*wh*, 'For every person, who invited that person?'

b. *OSV (scrambled) [QP-object... Wh-subject...] question:*
Daremo$_i$-o dare-ga t_i syootaisimasita ka.
everyone-ACC who-NOM invited Q
'Who invited everyone? (*scrambled*)'
Interpretation: *wh*>QP, 'For which person did that person invite everyone?'
*QP>*wh*, 'For every person, who invited that person?'

[11] Throughout this dissertation, the notation '*X* > *Y*' indicates that '*X* takes scope over *Y*', and '**X* > *Y*' indicates that the reading in which *X* takes scope over *Y* is not possible.
[12] A further example of inverse object-wide scope in Japanese occurs with unaccusative verbs. The locative object of an unaccusative verb can take inverse object-wide scope over the theme subject, as exemplified in (i) (Yatsushiro 1996: 324 (8b)):

i. Daremo-ga dokoka-ni tuita. *Interpretation*: S>O-LOC; O-LOC>S
everyone-NOM somewhere-LOC arrived
'Everyone arrived somewhere.'

The present dissertation focuses on scope interpretation in transitive sentences. However, see Chapter 4 for pre-experimental investigation of quantification in unaccusative sentences.
[13] 'Q' in the glosses indicates a question-marking particle.

c *SOV [QP-subject... Wh-object...] question:*
?? Daremo-ga nani-o kaimasita ka.
everyone-NOM what-ACC bought Q
No interpretation due to unacceptability

d. *OSV (scrambled) [Wh-object... QP-subject...] question:*
Nani$_i$-o daremo-ga t_i kaimasita ka.
what-ACC everyone-NOM bought Q
'What did everyone buy? (*scrambled*)'
Interpretation: *wh*>QP, 'For what *x* did every person buy *x*?'
*QP>*wh*, 'For every person, what did that person buy?'

The non-scrambled and scrambled [*wh*-subject ... QP-object ...] questions, (18a) and (18b), allow only the reading in which the *wh*-subject *dare* 'who' takes wide scope, yielding the interpretation, 'For which person did that person invite everyone?'. Hence these questions can receive only an 'individual' answer such as *Hanako invited everyone*, and not a 'pair-list' answer such as *Hanako invited Jane, Yoshio invited Peter, etc.*, since the latter requires a QP-object-wide interpretation of the question. The non-scrambled [QP-subject... *wh*-object...] question (18c) is claimed to be unacceptable with a distributive interpretation of *daremo* 'everyone', and hence has no scope-dependent interpretation.[14] However, the scrambled variant in (18d) is acceptable, again with a *wh*-wide scope interpretation, like (18a–b). Thus (18d) allows an individual answer such as *Everyone bought red wine*, but not a pair-list answer such as *Hanako bought red wine, Yoshio bought beer, etc.* The relevant generalisation can be presented schematically as follows (Hoji 1985: 336, (54b–e)):

19. a. wh-NOM QP-ACC V. unambiguous: *wh*>QP; *QP>*wh*
b. QP$_i$-ACC wh-NOM t_i V. unambiguous: *wh*>QP; *QP>*wh*
c. *QP-NOM wh-ACC V. unacceptable
d. wh$_i$-ACC QP-NOM t_i V. unambiguous: *wh*>QP; *QP>*wh*

[14] Hoji (1985: 270) points out that the unacceptability of (18c) is heightened when a distributive reading of *daremo* 'everyone' is forced by means of the floating quantifier *sorezore* 'each', as in (i):

i. *Daremo-ga sorezore nani-o kaimasita ka
everyone-NOM each what-ACC bought Q
'What did everyone (= each person) buy?'

The marginality of (18c) is due to the possibility of a 'group' reading of *daremo* 'everyone', whereby *daremo* refers to a group of people who bought something together. Such a reading is scope-independent. Note that with a referential subject instead of a QP-subject, (18c) would be perfectly acceptable. Explanations by Hoji (1985), Aoun & Li (1993) and Tomioka (2004) of the unacceptability of [QP-subject ... *wh*-object...] questions in Japanese are described in the following sections of this chapter. Note, however, that not all linguists agree that (18c) is unacceptable. Miyagawa (1997) and Yoshida (1993) (both cited in Hagstrom 1998: 159) as well as Kuno & Takami (2002) find canonically ordered questions such as (18c) acceptable.

The unambiguity of (19d) provides the contrast between Japanese and English referred to in (13). While English [*wh*-subject...QP-object...] sentences are argued to allow only the *wh*-wide interpretation, like Japanese, English [*wh*-object...QP-subject...] questions are ambiguous. This is exemplified in (20a–b):

20. a. Who bought everything? unambiguous: *wh*>QP
    b. What did everyone buy? ambiguous: *wh*>QP; QP>*wh*

Thus, (20a) allows only an individual answer, such as *Jane bought everything*, but (20b) allows both an individual answer (e.g., *Everyone bought red wine*) or a pair-list answer (e.g., *Jane bought red wine, Sam bought beer, ...*).

The research for this dissertation focuses on canonical and scrambled Japanese QP-QP sentences, and on scrambled Japanese [*wh*-object...QP-subject...] questions. The differences between Japanese and English presented thus far, with respect to these structures, are summarised in Table 1.

Table 1: Scope interpretations in Japanese and English for different doubly-quantified structures (✓= 'available', ✗ = 'not available')[15]

Structure	Scope	Japanese	English
QP-subj...QP-obj...	S>O	✓	✓
	O>S	✗	✓
QP-obj$_i$...QP-subj...t_i	S>O	✓	*n/a*
	O>S	✓	*n/a*
Wh-obj...QP-subj...?[a]	*wh*>QP	✓	✓
	QP>*wh*	✗	✓

a. This structure is scrambled in Japanese, but not in English.

The discussion turns next to Chinese and Korean. Like Japanese, both of these languages are claimed to exhibit 'scope rigidity' (see, e.g., Aoun & Li 1989, 1993; Huang 1982 for Chinese; see Beck & Kim 1997; Kim 1989; Kim & Larson

[15] The mechanism of scrambling is assumed not to apply in English. Although topicalisation of an object is possible in English, as in (i), it is highly marked and requires a pause, as indicated by the comma after *languages*.

i. Two languages, everyone speaks.

Scrambling in Japanese is far less marked than the topicalisation in (i), and scrambled elements do not need to be followed by a pause. Reinhart (1978) discusses sentences such as (i) with reference to scope interpretation. They are argued to be ambiguous, allowing both subject-wide and object-wide scope.

1989; Sohn 1995 for Korean).[16] In other words, they also do not allow inverse scope interpretations in doubly-quantified sentences with canonical word order. Examples are presented in (21) for SVO Chinese and (22) for SOV Korean:[17]

21. Meige xuesheng dou mai-le yiben shu — unambiguous: S>O; *O>S
    every student all buy-ASP one book
    'Every student bought a book'
    (Huang 1982: 112, (3))

22. Nwukwunka-ka nwukwuna-lul chodayhatta. — unambiguous: S>O; *O>S
    someone-NOM everyone-ACC invited
    'Someone invited everyone.'
    (Kim 1989: 366, (24a))

Scrambling is not available in Chinese, but in Korean, the object can be scrambled over the subject just as in Japanese, resulting in an ambiguous sentence with both the subject-wide and object-wide interpretations:

23. Nwukwuna$_i$-lul nwukwuna-ka t_i chodayhatta. — ambiguous: S>O; O>S
    everyone-ACC someone-NOM invited
    'Someone invited everyone. (*scrambled*)'
    (Kim 1989: 367, (28a))

Regarding [QP-subject... *wh*-object ...] questions, Chinese appears to pattern with English in allowing both individual (*wh*-object-wide) and pair-list (QP-subject-wide) readings:[18]

24. Meigeren dou maile shenme? — ambiguous: *wh*>QP; QP>*wh*
    everyone all bought what
    'What did everyone buy?'
    (based on Aoun & Li 1993: 59, (42))

Korean, on the other hand, is claimed to lack pair-list readings.[19] Kim (1989) asserts that Korean questions with a QP-subject and *wh*-object behave in the same way as the equivalent Japanese sentences. With non-scrambled word order, the questions are

---

[16] Kim & Larson (1989) note that object-wide scope is available in a Korean SOV QP-QP sentence if the verb is a psych-predicate such as *hwuhoyslepta* 'be regrettable'.
[17] The experimental part of this dissertation investigates doubly-quantified sentences with an existential subject and universally quantified object. Hence, where possible, examples with this structure are used. However, in Chinese, indefinite NPs are often (but not always: see Liu 1997) somewhat unnatural in sentence-initial subject position. Consequently, Chinese examples in the quantification literature—and hence also in this chapter—usually have a universally quantified subject and existentially quantified object, to avoid the potentially unnatural indefinite subject.
[18] Since Chinese is a *wh*-in situ language and it does not have scrambling, it does not have the [*wh*-object...QP-subject...] word order discussed for English and Japanese.
[19] This claim is challenged in Chapter 5, based on the quantitative experimental data collected from native speakers of Korean for this dissertation. However, for the purpose of the present chapter, the claim that Korean lacks pair-list readings is assumed to hold.

unacceptable, as shown in (25a) (*cf.* Japanese (18c)), while the scrambled question is claimed to be acceptable with an individual reading (*wh*>QP) but not a pair-list reading (QP>*wh*), as illustrated in (25b) (*cf.* Japanese (18d)).

25. a. ?* Nwukwuna-ka nwukwu$_i$-lul chodayhat ni.
everyone-NOM who-ACC invited Q (Kim 1989: 364, (18d))

b. Nwukwu$_i$-lul nwukwuna-ka t_i chodayhat ni. *wh*>QP; *QP>*wh*
who-ACC everyone-NOM invited Q
'Who did everyone invite? (*scrambled*)' (Kim 1989: 367, (29c))

Table 2 summarises the scope interpretation patterns attested in the literature for the languages and the structures under consideration.

Table 2: Scope interpretation in Japanese, English, Chinese and Korean for different doubly-quantified structures (✓= 'available', × = 'not available')

Structure	Scope	Japanese	English	Chinese	Korean
QP-subj...QP-obj...	S>O	✓	✓	✓	✓
	O>S	×	✓	×	×
QP-obj$_i$...QP-subj...t_i	S>O	✓	*n/a*	*n/a*	✓
	O>S	✓	*n/a*	*n/a*	✓
Wh-obj...QP-subj...?[a]	*wh*>QP	✓	✓	✓	✓
	QP>*wh*	×	✓	✓	×

a. This structure is scrambled in Japanese and Korean, but not in English and Chinese. In Chinese the word order is actually [QP-subj... *wh*-obj...].

It should be noted at this point that the generalisations summarised in Table 2 are not always agreed upon by all speakers. Judgements about scope interpretation are subtle and subject to individual variation.[20] Quantitative empirical data help to clarify the situation. As Chapter 3 will show, empirical investigations of native English (e.g., Kurtzman & MacDonald 1993; Lee, Yip & Wang 1999a, 1999b) and native Chinese (Lee, Yip & Wang 1999b) largely confirm the claims of scope ambiguity for English active QP-QP sentences and scope rigidity for Chinese active QP-QP sentences. However, quantitative empirical data are sparse or non-existent on Japanese and Korean QP-QP sentences, and on [*wh*-object...QP-subject...] questions in all four languages under consideration. The only relevant studies, to the author's

[20] Among others, Kuroda (1970), Reinhart (1997), and Tomioka (2004) comment on the vagueness of judgements about scope interpretation, or the variability among native speakers.

knowledge, are Sano (2004), who reports that 10 adult native-Japanese speakers made judgements generally as predicted by Hoji (1985) about subject-wide and object-wide interpretations of the SOV and OSV variants of one Japanese QP-QP sentence; and Miyamoto & Yamane (1996), who showed that five native speakers of English accepted both individual and pair-list interpretations of English [*wh*-object...QP-subject...] questions (details of these studies are presented in Chapter 3). Thus, one contribution of the present study is to fill some gaps in our cross-linguistic database on quantifier scope interpretation by conducting the relevant quantitative research. The results of this research are reported in Chapters 4 and 5. However, for the purpose of the present chapter, the judgements summarised in Table 2 are assumed to hold, and individual variation is ignored, unless indicated otherwise.

Two poverty-of-the-stimulus problems for the L2 acquisition of scope interpretation phenomena in Japanese are instantiated in the data summarised in Table 2. They are (i) knowledge of the lack of object-wide (O>S) scope in Japanese by learners whose L1 is English, and (ii) knowledge of the lack of pair-list reading in (scrambled) Japanese [*wh*-object ... QP-subject...] questions by learners whose L1 is English or Chinese. I assume (following Full Transfer/Full Access, outlined in the previous chapter) that L1 knowledge of scope interpretation is transferred to the interlanguage at the start of L2 acquisition. Thus, the interlanguage of English learners of Japanese will initially allow inverse scope in non-scrambled Japanese QP-QP sentences and pair-list readings in scrambled Japanese [*wh*-object ... QP-subject...] questions, due to transfer from English; and the interlanguage of Chinese learners of Japanese will initially allow pair-list readings (but not inverse scope in non-scrambled QP-QP sentences), due to transfer from Chinese. The interlanguage of Korean learners of Japanese will be like Japanese from the outset, with respect to the scope interpretation phenomena in Table 2, due to transfer from Korean. In order for the interlanguage of English and Chinese learners of Japanese to become native-like, it must be restructured so that English learners come to know that Japanese lacks object-wide scope in SOV sentences, and both English and Chinese learners come to know that Japanese lacks pair-list interpretations of scrambled [*wh*-object...QP-subject...] questions. These are poverty-of-the-stimulus problems, because there can be no direct evidence in the target language input for the lack of phenomena which are already instantiated in the interlanguage due to L1 transfer. Even though learners never encounter Japanese SOV QP-QP sentences in an unambiguous object-wide scope context or scrambled [*wh*-object...QP-subject...] questions in an unambiguous pair-list context, failure to encounter such interpretations does not logically preclude their existence, given that the interlanguage allows for them to exist. Thus the target language input can contain no direct motivation for the relevant restructuring of the L1 English-L2 Japanese and L1 Chinese-L2 Japanese interlanguage.

This dissertation investigates whether these two poverty-of-the-stimulus problems can nonetheless be overcome by English-speaking and Chinese-speaking learners of Japanese. As pointed out in the previous chapter, a poverty-of-the-stimulus problem is a poverty-of-the-stimulus problem whatever the theory that accounts for it. However, different theories make different predictions about how UG enables poverty-of-the-stimulus problems to be overcome, and, in the case of L2 poverty of the stimulus, about whether the problems can be overcome at all. Sections 2.2 and 2.3, following, outline a number of theoretical accounts of the scope interpretation patterns summarised in Table 2, and comment on what each account entails for L2 acquisition of the Japanese patterns by learners whose L1s are English, Korean or Chinese.

2.2. 'All quantifiers are equal': analyses focusing on *every* and *some*

This section presents a summary of four syntactic accounts of quantifier scope assignment which focus predominantly on the universal quantifier *every* and the existential *some*. The seminal works of May (1977, 1985) are presented first, since, although these address only English, they are the foundation on which many subsequent accounts are based. This is followed by an overview of Hoji (1985), the first major study of scope interpretation differences between English and Japanese. Aoun & Li's (1989, 1993) comprehensive account of scope in English, Chinese and Japanese is presented next. These first three accounts are all set within the *Government and Binding* (Chomsky 1981)/*Barriers* (Chomsky 1986) framework of syntactic theory. The final section presents an outline of Hornstein's (1995) account, set within the Minimalist framework.

2.2.1. Quantifier Raising and Logical Form: May 1977, 1985

The question of how to account for quantification phenomena has animated philosophers and mathematicians since (at least) medieval times (Heim & Kratzer

1998). The problem is that a quantifier in object position can take scope over the subject that precedes it, as indicated in (26).

26. Jane saw everyone.
 Interpretation: $\forall x$ saw(Jane,x) (= For every x where x is a person, Jane saw x)

Attempting to account for such non-linear interpretation facts in generative grammar led May (1977) to propose the existence of an abstract level of linguistic representation, namely Logical Form (LF). He suggests that quantifiers undergo obligatory Quantifier Raising (QR)—covert movement at LF—to the position in which they take scope. Thus (26) has the following LF representation:

27. [$_S$ everyone$_i$ [$_S$ Jane saw t_i]]

Everyone in (27) has moved, leaving a bound variable trace, so that it adjoins to, and thus c-commands, the sentence it takes scope over. The c-command relationship is integral to May's (1977: 11) definition of scope:[21]

28. The scope of a quantified phrase α is everything which it c-commands [at LF].

For ambiguous sentences containing two QPs, two representations are required: one in which QP1 c-commands QP2, and one in which QP2 c-commands QP1. May proposes that the application of QR is not ordered, and thus the two relevant LF structures are both possible, as in (29), where (29b) and (29c) represent the two LF representations of (29a).[22]

29. a. Every student admires some professor.
 b. [$_{IP}$ every student$_i$ [$_{IP}$ some professor$_j$ [$_{IP}$ t_i admires t_j]]]
 c. [$_{IP}$ some professor$_j$ [$_{IP}$ every student$_i$ [$_{IP}$ t_i admires t_j]]]

In (29b), *every student* c-commands *some professor*, thus representing the interpretation whereby 'for every student there is some professor whom she/he

---

[21] May uses Reinhart's (1976) definition of c-command:

i. α c-commands β iff the first branching node dominating α dominates β, and α does not dominate β.

[22] In May (1977, 1985) and also Hoji (1985) 'S' and 'S-bar' are used for the projections now familiarly denoted by 'IP' and 'CP'. Henceforth, these projections are rendered as 'IP' and 'CP' in the relevant examples.

admires'. In (29c), *some professor* c-commands *every student*, thus capturing the inverse scope interpretation of (29a): 'there is some professor who is admired by every student'.

May (1985) reworks and extends May (1977). The representation in (29b) is abandoned as a possible LF representation due to the Empty Category Principle (ECP). The ECP, as it is instantiated in Government and Binding Theory (Chomsky 1981), requires that traces are 'properly governed'. This renders (29b) an invalid representation because the object *some professor* intervenes between the subject *every student* and its trace, blocking proper government of the trace.[23] In (29c), *every student* is adjacent to the subject trace, and thus can properly antecedent-govern it. (The object trace in (29b) and (29c) is assumed to be lexically governed by the verb.) May (1985) thus proposes that (29c) is the only licit LF representation of (29a). This representation is argued nonetheless to represent the ambiguity of (29a) by appealing to Aoun & Sportiche's (1983) maximal projection definition of c-command:

30. α c-commands β iff every maximal projection dominating α dominates β, and α does not dominate β.

Crucial to May's theory is a stipulation (also following Aoun & Sportiche 1983) that IP ('S') is a non-maximal projection with CP ('S-bar') as its maximal projection (May 1985: 34). With this stipulation and the definition in (30), *some professor* and *every student* c-command each other in (29c). Thus (29c) is compatible with both the subject-wide and object-wide scope interpretations.

May (1985) goes on to explore the interaction of *wh*-words and quantifiers, not addressed in May (1977). As noted in Section 1, English [*wh*-subject ... QP-object] questions are unambiguous, while [*wh*-object ... QP-subject] questions are ambiguous. For the *wh*-object questions, such as *What did everyone buy for Max?*, the LF structure in (31) is proposed (May 1985: 38 (14)):

31. [$_{CP}$ what$_j$ [$_{IP}$ everyone$_i$ [$_{IP}$ t_i bought t_j for Max]]]

---

[23] May (1985: 32) states that 'proper government' for subjects entails an empty category being '"closely" bound by a phrase in an A'-position'. He draws an analogy between the proposed ECP-based invalidity of (29b) and the ungrammaticality, due to that-trace violation, of (i). In (i), the intervention of *that* between the subject trace and its antecedent is argued to violate the ECP and thus account for the ungrammaticality (May 1985: 32, (2)):

i. *[$_{CP}$ who$_i$ do you believe [$_{CP}$ t_i that [$_{IP}$ t_i suspected Philby]]]

What and *everyone* c-command each other in (31) (using the maximal projection definition of c-command given in (30)). Thus either can take scope over the other, yielding the two interpretations of the question. However, for *wh*-subject questions, such as *Who bought everything for Max?*, an ambiguous LF representation parallel to (31) violates the ECP (May 1985: 41, (21)):

32. [CP who$_i$ [IP everything$_j$ [IP t_i bought t_j for Max]]]

The subject trace in (32) is not properly governed due to the intervening QP blocking government. In order to represent the unambiguity of [*wh*-subject ... QP-object] questions in English, May (1985: 41–42) proposes that QR is not limited to adjunction to IP: QPs can be adjoined to other nodes as long as there is no violation of relevant constraints (e.g., the ECP). In *Who bought everything for Max?*, VP is proposed as the adjunction site for QR (May 1985: 42, (22)):

33. [CP who$_i$ [IP t_i [VP everything$_j$ [VP bought t_j for Max]]]]

In (33), the ECP is satisfied, and the correct scope reading obtains: *who* c-commands *everything*, but *everything* does not c-command *who*; thus *who* can take scope over *everything*, but not vice versa.

2.2.2. Applying QR to Japanese: Hoji 1985

A shortfall of May's analyses of QP-QP and *wh*-QP interaction is that they are problematic for Japanese, since they allow more scope ambiguity than Japanese exhibits. Recall that a doubly-quantified SOV sentence in Japanese is unambiguous:

34. QP-NOM QP-ACC V. unambiguous: S>O; *O>S

Hoji (1985) expands May's (1977) framework so that it can also account for Japanese.[24] In addition to QR, he requires the following condition and assumptions:

35. *Condition on scope interpretation in Japanese* (Hoji 1985: 248, (76))
At LF *QP$_i$ QP$_j$ t_j t_i where each member c-commands the member to its right.

[24] Since Hoji (1985) uses May's (1977) framework, his theory does not address the issue of the ECP. The question of ECP violations is thus ignored in the presentation of Hoji (1985).

36. *Assumptions*
    a. Move α leaves a trace optionally.
    b. QR does not adjoin to CP.

The condition on scope interpretation in Japanese (35) requires that surface c-command relationships must be maintained at LF.[25] In other words, it prohibits the LF representations proposed in May (1977) for object-wide scope (29c), in which the object moves at LF to c-command the subject which it was c-commanded by in the surface structure. The condition thus rules out the inverse scope reading of a Japanese [QP-NOM QP-ACC V] sentence as follows (based on Hoji 1985: 244):

37. a. *Surface structure*: [IP QP_i-NOM[VP QP_j-ACC V]]
    b. *LF S>O representation*: [IP QP_i-NOM [IP QP_j-ACC [IP t_i [VP t_j V]]]]
    c. *LF O>S representation*: *[IP QP_j-ACC [IP QP_i-NOM [IP t_i [VP t_j V]]]]

In (37b), the order of the two QPs is the same as the order of their respective traces, so the representation is valid under Hoji's condition on scope interpretation in Japanese, given in (35). In (37c), on the other hand, the order of the QPs is the reverse of the order of their respective traces, violating the condition on scope interpretation. Thus, the S>O reading of (37a) is possible and the O>S reading is not.

The assumption in (36a), 'Move α leaves a trace optionally' (following Lasnik & Saito 1984), is used to account for the ambiguity of scrambled doubly-quantified sentences in Japanese. Hoji (1985) argues that a sentence-initial scrambled object is adjoined to IP, leaving a trace in VP, as in (38) (Hoji 1985: 250 (85)):

38. [IP QP_i-ACC [IP QP-NOM [VP t_i V]]]

Applying May's (1977) QR, the scrambled form in (38) could have the LF representations in (39a) or (39b) (Hoji 1985: 251 (86a) & (86b)).

---

[25] As Hoji (1985: 248) notes, his scope condition is a reformulation of Huang's (1982: 220, (70)) 'General Condition on Scope Interpretation':

i. *General Condition on Scope Interpretation*
   Suppose A and B are both QPs or both Q-NPs or Q-expressions, then if A c-commands B at SS [Surface Structure], A also c-commands B at LF.

The accounts differ in that Hoji's condition (35) is Japanese-specific, whereas Huang proposed the condition in (i) as a linguistic primitive and posited an English-specific restructuring rule to account for inverse scope in English. The restructuring rule vacuously moves an object-QP rightwards to adjoin to IP prior to QR, so that the correct c-command relationship (i.e., the object c-commanding the subject) is established before LF movement (Huang 1982: 232).

39. a. [IP QP$_i$-ACC [IP QP$_j$-NOM [IP t_i [IP t_j [VP t_i V]]]]]
b. [IP QP$_j$-NOM [IP QP$_i$-ACC [IP t_i [IP t_j [VP t_i V]]]]]

(39a) yields the O>S interpretation. In (39b), the correct c-command relationship holds between subject and object for the S>O interpretation, but this structure violates Hoji's condition on scope interpretation, since the ungrammatical [QP$_i$ QP$_j$ t_j t_i ...] order obtains. However, if it is assumed that Move α leaves a trace only optionally (as 36a), the intermediate subject trace in (39b) need not be present. The S>O interpretation can thus be derived from the LF representation in (40), in which the scope interpretation condition is upheld (Hoji 1985: 252, (87)):

40. [IP QP$_j$-NOM [IP QP$_i$-ACC [IP _ [IP t_j [VP t_i V]]]]]

Hoji's second assumption (36b), that QR does not adjoin to CP, is required for his account of *wh*-QP interaction in Japanese. The assumption accounts for Hoji's observation (described in Section 1) that the *wh*-word in Japanese always has scope over the QP. Hoji assumes that *wh*-words in Japanese move to CP at LF (although they are in situ in the surface order). Thus, if QR allowed quantifiers to adjoin to CP, they would c-command a *wh*-word in CP, and presumably take scope over it. The combination of the condition on scope interpretation in Japanese (35) and the assumptions in (36a) and (36b) correctly yield a single LF representation for each of the *wh*-subject and *wh*-object questions forms given in (18) and repeated below in (41) with the LF representations proposed by Hoji (1985: 264, 267).

41. a. *SOV [Wh-subject... QP-object...] question:*
Dare-ga daremo-o syootaisimasita ka. *wh*>QP; *QP>*wh*
who-NOM everyone-ACC invited Q
'Who invited everyone?'
LF: [CP dare$_j$-ga [IP daremo$_i$-o [IP t_j [VP t_i syootaisimasita]]] ka]

b. *OSV (scrambled) [QP-object... Wh-subject...] question:*
Daremo$_i$-o dare-ga t_i syootaisimasita ka. *wh*>QP; *QP>*wh*
everyone-ACC who-NOM invited Q
'Who invited everyone? (*scrambled*)'
LF: [CP dare$_j$-ga[IP daremo$_i$-o [IP _ [IP t_j [VP t_i syootaisimasita]]]] ka]

c. *SOV [QP-subject... Wh-object...] question:*
?? Daremo-ga nani-o kaimasita ka. unacceptable
everyone-NOM what-ACC bought Q
LF: *[CP nani$_i$-o [IP daremo$_j$-ga [IP t_j [VP t_i syootaisimasita]]] ka]

d. *OSV (scrambled) [Wh-object... QP-subject...] question:*
Nani$_i$-o daremo-ga t_i kaimasita ka. *wh*>QP; *QP>*wh*
what-ACC everyone-NOM bought Q
'What did everyone buy? (*scrambled*)'
LF: [$_{CP}$ nani$_i$-o [$_{IP}$ daremo$_j$-ga [$_{IP}$ t_i [$_{IP}$ t_j [$_{VP}$ t_i syootaisimasita]]]] ka]

The order of the *wh*-word and the QP in each LF representation in (41) matches the order of their respective traces, except in (41c). In other words, the LF for (41c) violates Hoji's condition on scope interpretation. This, Hoji (1985: 264) argues, accounts for the ungrammaticality of questions like (41c). The assumption that QPs cannot adjoin to CP rules out repair of the violation by QR of the subject above the *wh*-object (and concomitant intermediate trace deletion). The unambiguity of the three grammatical question forms, (41a), (41b) and (41d), is similarly assured by the assumptions. (41a) has only one possible LF representation since IP is assumed to be the only possible adjunction site for the QP. (41b) could theoretically have another LF representation in which the intermediate trace optionally remained, but this would yield a trace c-command relationship that differed from the QP c-command relationship and hence violate the condition on scope interpretation (35). Similarly, (41d) could have another representation in which the intermediate trace is optionally deleted, but again this would violate Hoji's condition on scope interpretation.

2.2.2.1. Comment

Hoji (1985) accounts for Japanese QP-QP and *wh*-QP interactions using May's (1977) model and some additional assumptions. He notes, however, that his framework cannot account for the English ambiguity in questions like *What did everyone buy?* without a language-specific stipulation that QR in English can adjoin QPs to CP (Hoji 1985: 338). Explanation of the ambiguity of English [*wh*-object...QP-subject...] questions is precisely the problem that May (1985) attempts to solve (as detailed in Section 2.2.1, above). However, May (1985) makes false predictions of ambiguity for some of the Japanese questions in (41). These problems highlight the difficulty of accounting for the cross-linguistic differences in scope interpretation.

From the point of view of L2 acquisition of Japanese scope rigidity, the key element of Hoji's framework is the Japanese-specific condition on scope interpretation. This condition could presumably be recast as a UG parameter on scope interpretation with an 'unambiguous scope' setting and an 'ambiguous scope' setting. In L1 acquisition, the 'unambiguous' setting could be a default which changes to 'ambiguous' when positive evidence for scope ambiguity is encountered in the input. Thus, in native Japanese, Korean and Chinese, the parameter remains set at 'unambiguous' while in English it is reset to 'ambiguous'.[26] Assuming that L1 parameter settings transfer to the interlanguage (in accordance with Full Transfer/Full Access, described in Chapter 1), Chinese and Korean learners of Japanese will have knowledge from the outset of Japanese scope rigidity in SOV QP-QP sentences due to the 'unambiguous' setting of their scope interpretation parameter. English learners of Japanese, on the other hand, will have to reset this parameter. Since there can be no direct evidence for the lack of inverse scope in Japanese, it seems logically impossible for such a parameter to be reset. Thus, under this account English learners of Japanese are predicted to be unable to acquire Japanese scope rigidity in SOV QP-QP sentences.

However, it could be the case that the cross-linguistic differences between rigid scope languages and ambiguous scope languages do not derive from a dedicated scope interpretation parameter. They could, instead, be the consequences of some other feature or features of the relevant languages. This is indeed the claim in the analyses by Aoun & Li (1993) and Hornstein (1995), presented in the next two sections. Both these analyses make use of a further development in syntactic theory to account for cross-linguistic differences in scope interpretation: the VP-internal subject hypothesis.

[26] As noted above, further stipulations or mechanisms would be required under Hoji (1985) to account for the ambiguity of English [*wh*-object ... QP-subject...] questions, and presumably also for the ambiguity of Chinese [QP-subject ... *wh*-object ...] questions. The remainder of the discussion, following, focuses on acquisition of scope interpretation in [QP... QP...] sentences, and does not make predictions for [*wh*-object...QP-subject...] questions under Hoji (1985).

2.2.3. A universal account of cross-linguistic scope interpretation phenomena: Aoun & Li (1993)[27]

A key feature of Aoun & Li's (1993) analysis of quantifier scope (henceforth 'A&L') is that, in contrast to Hoji (1985) and Huang (1982), it aims to offer a universal account of cross-linguistic differences in quantifier interpretation (specifically between Chinese, Japanese and English), rather than positing language-specific rules.[28] A&L follow May (1977, 1985) in assuming that quantifiers are raised at LF and adjoined to an A'-position, leaving a variable in their base position. They then propose three universal principles as the key elements of their analysis:[29]

42. *The Minimal Binding Requirement* (MBR) (A&L: 19 (19))
    Variables must be bound by the most local potential antecedent (A'-binder).

43. *The Scope Principle* (A&L: 88 (50))
    An operator A may have scope over an operator B iff A c-commands B or an A'-element co-indexed with B.[30]

44. *Optionality of QR* (based on A&L: 83)
    QR applies optionally to a QP in a non-thematic position (θ'-position) and obligatorily to a QP in a thematic position (θ-position).[31]

In addition, A&L adopt the VP-internal subject hypothesis proposed by (among others) Kitagawa (1986), Koopman & Sportiche (1991) and Kuroda (1988). According to this hypothesis, languages differ with regard to whether subjects

---

[27] Aoun & Li (1993) expands on and refines Aoun & Li (1989).

[28] It was suggested in the previous section that Hoji's (1985) Japanese-specific condition on scope interpretation could potentially be recast as a UG parameter. As will be seen, A&L avoid positing a dedicated scope interpretation parameter.

[29] The *Optionality of QR* principle (44) is not stated as a specific principle by A&L. It is identified (under the name 'Condition on QR') by Kuno, Takami & Wu (1999: 65 (8)) in their overview of A&L.

[30] A&L use Reinhart's (1983: 23) definition of c-command (A&L: 201, fn. 8):

i. Node A c(onstituent)-commands node B iff the branching node α_1 most immediately dominating A either dominates B or is immediately dominated by a node α_2 that dominates B, and α_2 is of the same category type as α_1.

[31] The obligatory application of QR to QPs in θ-positions is argued to come from the θ-Criterion (Chomsky 1981: 36):

i. Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument.

A&L argue that since QPs are not referential expressions, they cannot bear θ-roles and hence must move out of θ-positions (A&L 1993: 79). The remaining variable then bears the θ-role. A&L do not explain what motivates QR from a non-θ-position, in the cases where this occurs.

undergo raising to Spec,IP or not: in English subjects are base-generated in Spec,VP and raise to Spec,IP, while in Chinese and Japanese subject-raising is not available, so subjects remain in Spec,VP, or they are base-generated in Spec,IP.[32] This difference in constituent structure between Chinese and Japanese on the one hand, and English on the other, is used, in conjunction with the principles in (42–44), to account for the cross-linguistic scope interpretation phenomena set out in Section 2.1, as is shown in the following sub-sections.[33]

2.2.3.1. QP-QP scope interaction in Aoun & Li (1993)

The ambiguous English QP-QP sentence in (45) receives the two LF representations in (46), in A&L's framework.[34]

45. [IP Every policeman [VP1 t_i [VP2 saw a thief]]] ambiguous: S>O; O>S

46. a. [IP every policeman$_i$ [IP x_i [VP1 t_i [VP2 a thief$_j$ [VP2 saw x_j]]]]]
    b. [IP a thief$_j$ [IP every policeman$_i$ [VP1 t_i [VP2 saw x_j]]]]

(46a) represents the S>O reading. Both QPs undergo QR: the subject *every policeman* adjoins to IP, and the object *a thief* to VP2. The MBR (42) is satisfied, since both variables x_i and x_j are bound by their most local potential A'-binder (i.e. by the QPs with which they are co-indexed, and which are in A'-positions). Thus, since the MBR is met, and *every policeman* c-commands *a thief*, the Scope Principle (43) allows for the S>O interpretation. Note that the NP-trace of subject-raising, t_i, is not involved in scope determination. In (46b), only the object *a thief* undergoes QR; the subject *every policeman* remains in Spec,IP. This accords with the *Optionality of QR* principle (44): Spec,IP is a θ'-position (the subject receiving its θ-role in

[32] In fact, A&L do not state specifically where they assume the Chinese subject to originate, although a number of their representations of Chinese sentences show it in Spec,VP (A&L 26 (40), 75 (3), 76 (8)). However, Yusa (1995: 323) and Hornstein (1995: 164) assume that, for A&L, the subject in Chinese is generated in Spec,IP. This may come from Aoun & Li (1989: 153): 'a Chinese subject occurs either in Spec position of VP or in Spec position of I[P]'. In this case, Spec,IP is assumed to be a θ-position in Chinese (but not in English). For A&L, it is not actually crucial where the subject originates, but it is for Hornstein (1995) as will be seen in Section 2.2.4.
[33] In addition to the active QP-QP sentences and the *wh*-QP questions presented in Sections 2.2.3.1–2, following, A&L also apply their analysis to a variety of other structures containing QPs, including double-object constructions, passives, and complex NPs. Only the active QP-QP sentences and *wh*-QP questions are relevant to the present dissertation, so details of the other structure types are omitted.
[34] A&L's analysis relies on a distinction between traces of NP-movement and variables left by QR. Therefore, throughout the presentation of A&L, 't' is used to denote the trace left by NP-movement, and 'x' is used to indicate a variable. In addition, note that A&L use a VP-shell structure in which the subject is generated in Spec,VP1.

Spec,VP), so QR out of Spec,IP is not obligatory. Consequently, there is space for the object *a thief* to adjoin to IP. The MBR is again satisfied: *a thief* is the closest potential A'-binder for the variable x_j. Thus, by the Scope Principle, the O>S interpretation obtains.

For Chinese and Japanese, the lack of subject-raising leads to the lack of an object-wide scope reading, as shown in (47) for Chinese and (48) for Japanese.

47. a. [IP [VP1 Meige jingcha [VP2 dou kandaoyigexiaotou]]] S>O; *O>S
every policeman all see one thief
'Every policeman saw a thief.' (based on A&L: 14 (9b) and 75 (3))

LF representations:
b. [IP [VP1 meige jingcha$_i$ [VP1 x_i [VP2 yige xiaotou$_j$ [VP2 dou kandao x_j]]]]]
c. *[IP [VP1 yige xiaotou$_j$ [VP1 meige jingcha$_i$ [VP1 x_i [VP2 dou kandao x_j]]]]]

48. a. [IP [VP1 Dareka-ga [VP2 daremo-o semeta]]] S>O; *O>S
someone-NOM everyone-ACC criticised
'Someone criticised everyone.'

LF representations:
b. [IP [VP1 dareka-ga$_i$ [VP1 x_i [VP2 daremo-o$_j$ [VP2 x_j semeta]]]]]
c. *[IP [VP1 daremo-o$_j$ [VP1 dareka-ga$_i$ [VP1 x_i [VP2 x_j semeta]]]]]

Both the subject and object in (47) and (48) are in θ-positions and consequently, by the *Optionality of QR* principle (44), must undergo QR. This yields the licit LF structures in (47b) and (48b) in which the subject adjoins to VP1 and the object to VP2.[35] The MBR is satisfied, since each QP is the most local potential A'-binder for its respective variable. The subject in each case c-commands, and thus, by the Scope Principle, takes scope over the object, yielding the linear S>O scope relation attested in the literature. The LF representations in (47c) and (48c), in which the inverse scope relation would result (i.e., in which the object c-commands the subject), are ruled out by the MBR. In these representations the most local potential A'-binder for both variables is the subject. Since, in each case, the subject is co-indexed with its own variable and cannot also bind a variable bearing a different index, the object variable x_j, cannot be bound by its most local potential A'-binder, and the

[35] The subject in each case could equally adjoin to IP at LF. The resulting configuration would be the equivalent to adjunction to VP1 as far as scope relations are concerned. Similarly, in (47c) and (48c), the object could raise to IP, but it would result in the same violation of the MBR as in its VP1-adjoined position in (47c).

representation fails. Thus, the assumption, derived from the VP-internal subject hypothesis, that Chinese and Japanese subjects are in a θ-position prior to LF, while English subjects are in θ'-position, results in the lack of inverse scope in Chinese and Japanese.

A&L's account extends readily to scrambled QP-QP sentences in Japanese, which are ambiguous. A&L assume, following Kuroda (1988) that scrambling in Japanese is A-movement to Spec,IP. Since Spec,IP is not a θ-position, any QP scrambled to Spec,IP need not undergo QR. This results in licit LF representations for both S>O and O>S readings of the scrambled version of (48), as shown in (49).

49. a. [IP Daremo-o [VP1 dareka-ga [VP2 t_j semeta]]] S>O; O>S
everyone-ACC someone-NOM criticised
'Someone criticised everyone.' (*scrambled*)

LF representations:
b. [IP dareka-ga$_i$ [IP daremo-o$_j$ [VP1 x_i [VP2 [VP2 t_j semeta]]]]]
c. [IP daremo-o$_j$ [IP x_j [VP1 dareka-ga$_i$ [VP1 x_i [VP2 t_j semeta]]]]]

In (49b), the object *daremo* 'everyone' optionally does not undergo QR. The subject *dareka* 'someone' must undergo QR since it is in a θ-position. The MBR is satisfied, since *dareka* is the most local potential A'-binder for its own variable x_i (*daremo* not being in an A'-position and therefore not being a potential A'-binder). (49b) thus yields the subject-wide scope interpretation of (49a). The object-wide intepretation is obtained from (49c), in which *daremo* 'everyone' optionally undergoes QR to adjoin to IP, while *dareka* 'someone' adjoins to VP1. Again there is no violation of the MBR, so the O>S reading obtains. The ambiguity of scrambled Japanese QP-QP sentences is thus accounted for.

2.2.3.2. *Wh*-QP scope interaction in Aoun & Li (1993)

As described in Section 1, Chinese patterns with English, with regard to the interpretation of questions with a QP-subject and *wh*-object. This is exemplified in (50) (presented previously in (24)), which allows both an individual reading (*wh*>QP) and a pair-list reading (QP>*wh*).

50. Meigeren dou maile shenme? *wh*>QP; QP>*wh*
everyone all bought what
'What did everyone buy?' (based on A&L: 59 (42))

In Japanese, on the other hand, the pair-list reading is claimed to be lacking (Hoji 1985). On [*wh*-subject ... QP-object ...] questions, on the other hand, all three languages are unambiguous, allowing only the individual reading, as exemplified for Chinese in (51).

51. Shei maile meige dongxi? *who*>QP; *QP>*who*
who bought every thing
'Who bought everything?' (based on A&L: 7 (22))

A&L make use of a number of further assumptions to account for these scope interpretation phenomena. Key to whether or not a pair-list reading is available in [*wh*-object ... QP-subject...] questions are the assumptions that (i) *wh*-movement and QR are qualitatively different types of movement—*wh*-movement allows intermediate traces to be left while QR does not; and (ii) *wh*-words undergo *wh*-movement to Spec,CP in Chinese and English (covertly, at LF, in Chinese), while, in Japanese, *wh*-words undergo QR.

Considering the unambiguity in Chinese and English of [*wh*-subject ... QP-object] questions first, this is straightforwardly explained for (51) (and its English equivalent) by the LF representations in (51a–b) (A&L: 86).

52. a. [CP shei$_i$ [IP [VP1 x_i [VP2 meige dongxi$_j$ [VP2 maile x_j]]]]]
b. [CP who$_i$ [IP x_i [VP1 t_i [VP2 everything$_j$ [VP2 bought x_j]]]]]

In both (52a) (Chinese) and (52b) (English), the *wh*-subject c-commands the QP-object, and thus takes scope over it. The MBR is satisfied, since the variables left by *wh*-movement and QR are bound by their respective operators, which are their most local potential A'-binders. However, if the QP-object raised to adjoin to CP (the only adjunction site that would c-command *shei/who*), the MBR would be violated, as the QP-object would no longer be the most local potential A'-binder for its own variable. Thus the question in (51) has only one possible scope reading, the *wh*>QP reading which gives rise to an individual answer, such as *Jane did.*

The ambiguity of Chinese [QP-subject...*wh*-object ...] questions, such as (50), and English [*wh*-object...QP-subject...] questions, is more complex. A&L make use of the assumption referred to above, whereby, following Chomsky (1986), *wh*-operators have the possibility of leaving an intermediate trace on their way to Spec,CP (A&L: 86). In addition, they assume a distinction between variables bound by QPs and variables bound by *wh*-operators (A&L 56–57). A variable bound by a QP is not subject to Principle C (following Aoun & Hornstein 1985), while a variable bound by a *wh*-operator is (following Chomsky 1981).[36] Consequently, the LF representation schematised in (53) is possible, even though the parallel representation in which the two operators are QPs, as in (54), is ruled out (A&L: 68).

53. [CP wh$_j$ [IP/VP1 QP$_i$ [IP/VP1 x_i [VP2 V x_j]]]]

54. *[CP QP$_j$ [IP/VP1 QP$_i$ [IP/VP1 x_i [VP2 V x_j]]]]

The configuration in (53) does not violate the MBR, because Principle C ensures that QP$_i$ is not a potential A'-binder for the object variable x_j. If x_j—an R-expression—were bound by QP$_i$, it would end up being co-indexed with and also A-bound by x_i (since x_i c-commands x_j), in violation of Principle C. Thus the only potential A'-binder for x_j is *wh*$_j$. In (54), since x_j is not an R-expression (as a variable left by a QP), it is not subject to Principle C, and hence QP$_i$ is its most local potential A'-binder. (54) thus violates the MBR and is ruled out.

Bearing these assumptions in mind, the Chinese and English LF representations for (50) are as illustrated in (55), with the optional intermediate trace indicated in bold.

55. a. [CP shenme$_j$ [IP [VP1 meigeren$_i$ [VP1 x_i **[VP2 t_j** [VP2 dou maile x_j]]]]]]
b. [CP what$_j$ [C' did [IP everyone$_i$ [IP x_i [VP1 t_i **[VP2 t_j** [VP2 buy x_j]]]]]]]

Both (55a) and (55b) allow two possible scope readings, depending on the absence or presence of the optional intermediate trace. If the intermediate trace is absent, *shenme/what* c-commands *meigeren/everyone*, and the MBR is satisfied since the QP *meigeren/everyone* is not a potential A'-binder for the *wh*-variable x_j due to Principle C (as per (53)). Thus the *wh*-object>QP-subject scope relation obtains. If the intermediate trace is present, A&L's Scope Principle ((43), repeated in (56)) allows for it to play a role in determining scope relations.

56. *The Scope Principle* (A&L: 88, (50))
An operator A may have scope over an operator B iff A c-commands B or an A'-element co-indexed with B.

[36] Principle C of Binding Theory states (Chomsky 1986: 188) '[a]n R-expression [referring expression] is free'. Thus in A&L, the variable left by a *wh*-element is an R-expression, while the variable left by a QP is not.

The intermediate trace, t_i, in (55a–b) is an A'-element and it is co-indexed with *shenme/what*. *Meigeren/Everyone* c-commands this trace, the MBR is satisfied, and thus the QP-subject>*wh*-object scope obtains, in addition to the linear *wh*-object>QP-subject scope.

For Japanese, A&L address the full paradigm from Hoji (1985), presented in (19), and repeated below:

57. *Wh-QP interactions in Japanese*

a. wh-NOM QP-ACC V.	unambiguous: *wh*>QP; *QP>*wh*
b. QP$_i$-ACC wh-NOM t_i V.	unambiguous: *wh*>QP; *QP>*wh*
c. *QP-NOM wh-ACC V.	unacceptable
d. wh$_i$-ACC QP-NOM t_i V.	unambiguous: *wh*>QP; *QP>*wh*

As already noted, A&L assume that *wh*-words in Japanese do not undergo *wh*-movement. This assumption derives from Kim (1991) and Nishigauchi (1990), and is argued to be due to the obligatory Japanese question marker, *ka*, occurring as the head of CP and fulfilling the same role within the interrogative CP as *wh*-elements moved to Spec,CP in English or Chinese.[37] Specifically, the interrogative CP is argued to host an abstract [+wh] feature which must be satisfied at LF by a lexical item bearing a [+wh] feature within CP. This requirement is satisfied by *wh*-movement of *wh*-words to Spec,CP in English and Chinese (Aoun, Hornstein & Sportiche 1981; Rizzi 1991), while in Japanese the required [+wh] feature is borne by *ka* in the head of CP. Japanese *wh*-words thus do not need to undergo *wh*-movement. However, it is assumed that they do nonetheless move at LF, to a position governed by *ka*.[38] A&L (196) adopt the specific proposal of Kim (1991) that this type of movement is not *wh*-movement, but QR applied to *wh*-words.[39]

[37] Nishigauchi (1990: 18) notes that *ka* does not always mark questions in Japanese. In colloquial speech, the question marker could be *no* instead of *ka*, or the question could be indicated through intonation, without any overt question marker at all. Even in the latter case, the question intonation can be assumed to provide a covert form of *ka*, thus in all three cases, a question marker can be assumed to occupy C.

[38] A&L (203–4, fn. 2) assume the *Barriers* (Chomsky 1986) definition of government, following Aoun & Sportiche (1983: 214 (10)). The user-friendly version from Szabolcsi (2001: 617 (36)) is cited here:

i. α governs β = α c-commands β, and there are no maximal projection boundaries between them.

[39] A&L note that Nishigauchi (1990) also assumes the LF movement of *wh*-words in Japanese to a position governed by *ka*. However, Nishigauchi's proposal is that *wh*-elements move to Spec,CP and that this is indeed the same *wh*-movement undergone by *wh*-elements in languages without question markers, such as English. As will become clear, A&L's account requires the movement of Japanese *wh*-elements to be QR and not the *wh*-movement found in English and Chinese.

For A&L, '[t]he government requirement between the question marker *ka* and the *wh*-word [...] force[s] the *wh*-element to be adjoined to IP' (A&L: 196). This, in conjunction with the MBR and the other aspects of A&L's framework, results in only the LF representations which yield *wh*-wide scope for the questions in (57a, b and d), and additionally accounts for the ungrammaticality of (57c) (A&L: 196–197). The mechanism is outlined here just for (57d), the scrambled *wh*-object question, since this question form is the focus of investigation in the experimental part of this dissertation. The surface structure and LF form of (57d) are thus illustrated in (58a–b), respectively.

58. a. *Scrambled [wh-object ... QP-subject...] question*
[CP [IP Nani$_i$-o [VP1 daremo-ga [VP2 t_i kaimasita]]]C' ka]
what-ACC everyone-NOM bought Q
'What did everyone buy?' (*scrambled*)

b. [CP [IP nani$_j$-o [IP x_j [VP1 daremo$_i$-ga [VP1 x_i [VP2 t_j syootaisimasita]]]]]C ka]
(based on A&L: 196 (18))

(58b) is the only legitimate LF representation for (58a). The scrambled object *nani* 'what' undergoes QR to IP, in order to be governed by the question marker *ka* in the head of C.[40] The QP-subject *daremo* 'everyone' must undergo QR since it originates in the θ-position, Spec,VP1. The only possible adjunction site is VP1, since adjunction higher than *nani* would result in a structure like (59) which violates the MBR (with *nani*$_j$ becoming the most local potential A'-binder for x_i).

59. [QP$_i$ (=*daremo*, 'everyone') [QP$_j$ (=*nani*, 'what') [x_j [x_i [x_j V]]]]]

The distinction between *wh*-variables and QP-variables which legitimised the structure [wh$_j$ [QP$_i$ [x_i [V x_j]]]] in Chinese and English cannot apply here. A&L's claim is that, in Japanese, *wh*-words undergo QR, not *wh*-movement. Thus the variables left by the LF movement of Japanese *wh*-words are not *wh*-variables (i.e., variables bound by a *wh*-operator). This means that they are not subject to Principle C, so the possibility of x_i in (59) being bound by QP$_j$ cannot be ruled out in terms of a Principle C violation. (58b) thus yields the *wh*>QP scope relation, with the inverse relation being ruled out.

[40] Note that IP does not count as a barrier to government in Chomsky's (1986) *Barriers* framework. Thus the government relation given in footnote 38 obtains.

To summarise, the requirement observed by Hoji (1985) that Japanese *wh*-words must take wide scope (and that Japanese therefore lacks the pair-list interpretation of [*wh*-object ... QP-subject ...] questions), falls out from A&L's two-part assumption that (i) Japanese *wh*-words do not undergo *wh*-movement to Spec,CP, but (ii) they must undergo QR to IP in order to be governed by the question marker *ka* in C. A&L do not make claims about other languages; however, this account is clearly also applicable to Korean, which makes use of an obligatory sentence-final question marker *ni*, similar to the Japanese *ka*. For Chinese and English, the fact that *wh*-words can be under the scope of quantifiers in questions with a *wh*-object and QP-subject (hence allowing pair-list readings) is accounted for by the optional availability of an intermediate trace of *wh*-movement. Intermediate traces (crucially, in A'-position) can play a role in scope determination, so the LF structure in which the QP c-commands the *wh*-trace, [*wh*-object$_i$ [QP [...t_i ...]]], yields the QP>*wh* interpretation.

2.2.3.4. Comment

A&L's syntactic account of a broad range of scope phenomena in Chinese, English and Japanese is probably the most comprehensive cross-linguistic analysis of scope interpretation. Its strength is that it accounts for QP-QP interactions in terms of universal mechanisms, without resorting to language-specific principles. For *wh*-QP interactions, too, the assumptions accounting for the lack of pair-list readings in Japanese (i.e., that Japanese *wh*-words undergo QR to IP in order be governed by the question marker *ka*), could conceivably be restated in terms of a UG parameter concerning question particles. Such a parameter (or set of parameters) would have a 'question particle grammar' setting for languages like Japanese and Korean, one element of which would be the rule that within such a grammar the movement of *wh*-words is QR. The 'non-question particle grammar' setting for languages like English and Chinese, on the other hand, would entail that movement of *wh*-words is standard *wh*-movement. The lack of pair-list readings in question particle languages would then arise automatically from the correct setting of this UG option.[41]

In terms of the L2 acquisition of Japanese scope interpretation by native English speakers, A&L's account entails two acquisition problems: (i) the learners

[41] This hypothetical parameter would clearly require modification to account for any question-particle language found to allow pair-list interpretations in *wh*-QP questions.

must come to know that Japanese subjects—unlike English subjects—do not raise to IP; and (ii) they must come to know that LF movement of *wh*-words in Japanese is QR, not *wh*-movement as in English. The first problem is (hypothetically) solved by resetting the proposed VP-internal subject parameter, a corollary of which would be the ruling out of object-wide scope in SOV Japanese [QP... QP...] sentences; the second, by resetting the proposed question-particle parameter, which would rule out QP-wide scope in scrambled Japanese [*wh*-object...QP-subject...] questions. For Chinese-speaking learners of Japanese, only the second problem applies, while for Korean-speaking learners of Japanese neither problem applies, since their interlanguage grammar should be like the Japanese grammar from the outset.

Considering, first, the problem faced by both Chinese-speaking and English-speaking learners of acquiring knowledge that LF movement of Japanese *wh*-words is QR, this knowledge could arguably arise based on the ample evidence of question particles in the input. Since English and Chinese lack question particles, the L1-based English-Japanese and Chinese-Japanese interlanguage grammars will presumably fail to represent Japanese question particles and will consequently be forced by this failure to undergo restructuring to a grammar in which question particles are represented. In other words, internalisation of the evidence of question particles could activate a 'question particle grammar' option within UG. Thus, if A&L's theory is correct, English-speaking and Chinese-speaking learners of Japanese could readily overcome the poverty-of-the-stimulus problem represented by acquisition of the lack of pair-list readings in Japanese, due to the evidence of question particles in the input triggering the relevant restructuring of the interlanguage grammar.

The remaining problem for English-speaking learners of Japanese is resetting of the proposed VP-internal subject parameter. It is less obvious what a potential trigger for this could be. This is because there is a lack of clear empirical evidence to differentiate languages like Japanese, Chinese and Korean from languages like English with regard to the presence or absence of subject-raising. Consequently, there is disagreement about whether or not Japanese, Chinese and Korean really do lack subject-raising. Although Kitagawa (1986), Koopman & Sportiche (1991), Kuroda (1988) and Wible (1990) (among others) argue for the lack of subject-raising, Huang (1993), Ueda (1990), and others, present arguments that subject-raising does take place in these languages. The arguments for or against subject-raising in Japanese are usually concerned with the nature of INFL. The lack of agreement morphology in Japanese, for example, has led to suggestions that INFL is 'degenerate' in this language and consequently subjects do not raise to Spec,IP. Other arguments are theory-internal, centring around whether or not θ-roles and Case can be assigned in IP. It is not clear exactly how a VP-internal subject parameter might be set in L1 acquisition, but one possibility is that the 'default' setting is as proposed for Japanese: namely, the 'no-subject-raising' setting, where the subject is generated in the position it remains in. L1 acquirers of English might be able to reset this default to the 'subject-raising' option on the basis of evidence such as auxiliary verbs that intervene between the subject and the main verb.[42] However, once the parameter has been set to the 'subject-raising' option, it seems unlikely that it could be reset back to the default. This is because the L1-(English)-based interlanguage grammar can adequately—if, incorrectly—represent Japanese, as shown in (60a). The correct representation in A&L's theory, in which the subject remains in Spec,VP, is shown in (60b):

60. a. [$_{IP}$ [Dareka-ga]$_i$ [$_{VP}$ t_i [$_{V'}$ dono hon-mo [$_V$ yon-]]]$_I$ da]
    b. [$_{IP}$ [$_{VP}$ Dareka-ga [$_{V'}$ dono hon-mo [$_V$ yon-]]]$_I$ da]
    someone-NOM every book-QPt read- PAST
    'Someone read every book.'

It seems that there is no clear evidence that might motivate abandoning subject-raising. Thus, if it is impossible for English-speaking learners of Japanese to reset the proposed VP-internal subject parameter, then fossilization is predicted to occur with respect to this parameter. Consequently, they will not be able to able to acquire the lack of object-wide scope in Japanese SOV QP-QP sentences under this account.

The lack of consensus about whether languages really differ with respect to a VP-internal subject hypothesis is a potential threat to A&L's analysis. If it turns out that languages do not differ in this respect, the explanatory power of the analysis will be considerably compromised. A number of technical problems with the account have also been raised by Hornstein (1995), Kuno *et al.* (1999) and Yusa (1995) (see also Aoun & Li 2000; Kuno, Takami & Wu 2000). Finally, significant criticisms of the approach arise when it is viewed from the perspective of Minimalism (Chomsky

[42] Another possibility is that the 3rd person singular morphology of English triggers adoption of the 'subject-raising' option, although Sprouse (2000) argues convincingly against agreement morphology driving the acquisition of syntax.

1993, 1995). The *Optionality of QR* principle, for example, which is acceptable within the *Barriers* framework, is challenged in Minimalism by the question of why quantifiers in θ'-positions should ever raise, if there is no specific motivation for raising. In fact, as Hornstein (1995) describes, since all movement is argued to be feature-driven in the Minimalist framework, the very notion of QR is challenged, since it is movement for the purpose of scope assignment rather than for the purpose of checking features. Hornstein (1995) offers a Minimalist account of quantifier scope interpretation, as outlined in the next section.

2.2.4. Quantifier scope in Minimalism: Hornstein (1995)[43]

As noted in the previous section, the notion of QR becomes problematic with the adoption of Minimalist syntactic theory, since Minimalism requires that movement be feature driven. Thus *wh*-movement, for example, is motivated by the need for an abstract [+wh] feature on *wh*-words to be checked in Spec,CP. QR, which can adjoin quantifiers to any XP (i.e. does not have a specific landing site, in contrast to *wh*-movement), does not appear to have any (abstract) morphological motivation.

Hornstein's (1995) Minimalist account of quantifier scope appeals to the operation of feature-checking as the mechanism that leads to determination of scope relations; specifically, the checking of Case features. Subjects and objects within VP are assumed (following Chomsky 1993) to check off their respective nominative [+nom] and accusative [+acc] features in the functional subject and object agreement projections, AgrSP and AgrOP, above VP. In English, this happens by covert movement at LF from the VP-internal positions in which the subject and object are generated. Hornstein assumes that movement is a process of copying and deletion: the 'moved' element copies to its feature-checking position, then deletion applies at LF, so that one and only one member remains of the chain created by copying. For example, to check its [+nom] feature, a copy of the subject in VP is inserted into Spec,AgrSP, then either the subject in VP or the copy in Spec,AgrSP is deleted. Under this account, the following LF structure is obtained for the sentence *Someone attended every seminar*, prior to any deletion (Hornstein 1995: 155 (5)):

61. [$_{AgrSP}$ someone [$_{TP}$ [$_{AgrOP}$ every seminar [$_{VP}$ someone attended every seminar]]]]

[43] Kitahara's (1996) account of scope interpretation in English is very similar to Hornstein (1995). For space reasons, Kitahara's account is not presented in this dissertation.

One member of both the subject chain [someone ... someone ...] and the object chain [... every seminar ...every seminar] in (61) must be deleted. This allows for four possibilities, as shown in (62a–d), in which the deleted elements are indicated by curved brackets (Hornstein 1995: 155 (6a–6d)):

62. a. [$_{AgrSP}$ someone [$_{TP}$ [$_{AgrOP}$ every seminar [$_{VP}$ (someone) attended (every seminar)]]]]

b. [$_{AgrSP}$ someone [$_{TP}$ [$_{AgrOP}$ (every seminar) [$_{VP}$ (someone) attended every seminar]]]]

c. [$_{AgrSP}$ (someone) [$_{TP}$ [$_{AgrOP}$ (every seminar) [$_{VP}$ someone attended every seminar]]]]

d. [$_{AgrSP}$ (someone) [$_{TP}$ [$_{AgrOP}$ every seminar [$_{VP}$ someone attended (every seminar)]]]]

Of the four representations, (62b) and (62c) are ruled out by the assumption that, along the lines of Diesing's (1992) Mapping Principle, definite DPs/NPs must be in positions outside VP at LF. The strong quantifier *every* is considered to be definite, so its LF position cannot be inside VP, as it is in (62b–c). The remaining representations, (62a) and (62d), allow the two possible scope interpretations of the sentence. Hornstein assumes that scope is determined by the c-command relationship between the quantifiers themselves:

63. *Scope Principle* (Hornstein 1995: 154)
A quantified argument Q1 takes scope over a quantified argument Q2 iff Q1 c-commands Q2 [at LF].

Thus in (62a), *someone* takes scope over *every seminar* (subject-wide scope) and in (62d), *every seminar* takes scope over *someone* (object-wide scope).

To account for the unambiguity of QP-QP sentences in Chinese, Hornstein follows Aoun & Li (1993) in adopting the proposal that subjects do not raise in Chinese.[44] He assumes that subjects are generated in Spec.AgrSP.[45] Consequently, in a doubly-quantified SOV sentence in Chinese, there is no possibility of the object QP c-commanding a copy of the subject QP after raising to AgrOP for Case-checking,

---

[44] Recall from the preceding discussion (Section 2.2.3.4) that this assumption is not uncontroversial.
[45] See footnote 32. In Hornstein's framework, subjects in Chinese (and Japanese and Korean) cannot be generated in VP and remain there, since there would be no way of checking their [+nom] feature.

since AgrSP dominates AgrOP. The relevant LF structure, prior to deletion, is as follows (Hornstein 1995: 164 (37)):

64. [AgrSP subject [TP [AgrOP object [VP V object]]]]

Whichever copy of the object deletes, the remaining copy will be asymmetrically c-commanded by the subject, with the result that only the S>O scope reading obtains.

Unambiguity in Japanese non-scrambled doubly-quantified SOV sentences can be accounted for in the same way. The ambiguity of scrambled Japanese doubly-quantified sentences is accounted for as follows. Hornstein (1995: 94) adopts the position argued for in Saito (1992), that clause-internal scrambling is analysed at LF as A-movement to Spec,IP. In the articulated IP, Spec,AgrSP is a possible landing site of scrambling (presumably by allowing multiple specifiers, since the subject must also be in Spec,AgrSP). Thus the sentence in (65a) has the pre-deletion LF representation shown in (65b) (based on Hornstein 1995: 181 (105b), 182 (106)):

65. a. Dareka-o daremo-ga semeta.
someone-ACC everyone-NOM blamed
'Everyone blamed someone.'

b. [AgrSP dareka-o [AgrSP daremo-ga [TP [AgrOP dareka-o [VP dareka-o semeta]]]]]

Two of the three copies of *dareka-o* 'someone-ACC' must delete. If the copy in Spec,AgrSP remains, the O>S scope obtains. If this copy is deleted, the S>O scope obtains. Hence, the ambiguity of the scrambled Japanese sentence is accounted for.

Hornstein's (1995) treatment of *wh*-QP interactions is rather different. He follows Chierchia (1991) in taking the view that QPs do not quantify into questions. In other words, the presence or absence of a pair-list reading is not a function of *wh*-QP scope relations. Briefly, Chierchia (1991) considers the *wh*-element in a *wh*-QP question to consist of a function and an argument. The argument can be bound by a c-commanding quantifier, while the function is bound by the *wh*-element. Under this assumption, the pre-deletion LF structure of the [*wh*-object ... QP-subject ...] question in (66a) is represented as in (66b) (based on Hornstein 1995: 113 (55)):

66. a. What did everyone see?
b. [CP What$_i$ [AgrSP everyone$_j$ [TP [AgrOP what$_i$ [VP everyone$_j$ saw [*pro*$_j$ t_i]$_i$]]]]]

The [pro_j t_i]$_i$ structure in object position is a copy of *what* containing a covert pronoun argument bound by *everyone*. If that copy is not deleted (i.e., if *what* in Spec,CP and Spec,AgrOP are deleted), the question will have the pair-list interpretation. If that copy is among those deleted, the question will have the individual interpretation. In the case of [*wh*-subject...QP-object...] questions, such as (67a), the copy of *who* containing the covert pronoun must be among those deleted. If not, Hornstein (1995: 114) argues, a Weak Crossover (WCO) effect results because the pronoun is bound by a variable to its right (i.e., by *everyone* in object position within VP), as illustrated in (67b) (Hornstein 1995: 114 (57)):[46]

67. a. Who saw everyone?
    b. [CP Who$_i$ [AgrSP who$_i$ [TP [AgrOP everyone$_j$ [VP [pro_j t_i]$_i$ saw everyone$_j$]]]]]

In short, it is WCO, and not the nature of the *wh*-QP interaction that rules out the pair-list reading of [*wh*-subject...QP-object...] questions.

Since Chinese behaves in the same way as English with regard to *wh*-QP scope interpretations, Hornstein's analysis also accounts for Chinese. However, it is less obvious how the analysis could account for the lack of pair-list reading in scrambled Japanese [*wh*-object ... QP-subject ...] questions, such as that in (68).

68. Nani-o daremo-ga motte-kita no. unambiguous: *wh*>QP/*QP>*wh*
    What-ACC everyone-NOM brought Q
    'What did everyone bring?' (scrambled)

Applying Hornstein's analysis to (68) yields an LF representation as follows:

69. [CP nani-o$_i$ [AgrSP nani-o$_i$ [AgrSP daremo-ga$_j$ [TP [AgrOP nani-o$_i$ [VP [pro_j t_i]$_i$ motte-kita]]]]]C no]

The pair-list reading is not ruled out by this representation: since *daremo* precedes *pro*, a WCO effect does not arise. Thus, from the LF representation alone, the pair-list interpretation is predicted to be available. However, there is a further constraint on the availability of pair-list interpretations: the relevant quantifier must be of a type that can generate a set from which individuals can be mapped to a list. Not all

---

[46] Cf. other well-known WCO effects such as (i):

i. *His$_i$ mother loves everyone$_i$.

quantifiers are 'set generators'. Hornstein (1995: 113), following Chierchia (1991), points out that *most*, for example, in a question like *Who do most linguists admire?* does not support a pair-list answer. *Most* clearly differs from a universal quantifier like *everyone*, since the number of individuals indicated by *most* cannot have an upper or lower limit, in contrast to the number of individuals indicated by *everyone*, which must include every person in the context under consideration. It thus seems unlikely that a universal quantifier like *daremo* 'everyone' should fail to generate a set. However, the English universal *all* also lacks a pair-list reading, as indicated by the awkwardness of (70c) in answer to (70a).

70. a. Who do all kids love?
    b. Their mothers.
    c. *Jane loves Mickey Mouse, Sam loves Donald Duck, Tim loves Bugs Bunny,...

Although *all* in (70a) defines a set (i.e., the set that includes every child in the universe), (70c) indicates that individuals cannot be drawn from that set. In other words, the interpretation of *all* is collective, and does not distribute over a set. If *daremo* 'everyone' could be shown to share this property of the English universal *all*, the fact that it does not support a pair-list reading would not be incompatible with Hornstein's analysis. However, evidence indicates that *daremo* is characterised by distributivity. For example, Kawashima (1994: 136) claims that *daremo* in (71a) yields only a distributive interpretation, as shown in (71b), and not the collective interpretation in (71c).

71. a. Daremo-ga kooen de atta.
    everyone-NOM park in met
    b. 'Everyone met in the park.' (i.e., there were submeetings which each person engaged in at the park)
    c. *'The people all met in the park.'

It seems, therefore, that *daremo* 'everyone' meets the criterion of generating a set from which individuals can be drawn, and, as such, should a support a pair-list reading in [*wh*-object...QP-subject...] questions, under Hornstein's analysis. The fact that it does not is a problem for this account.

2.4.3. Comment

Hornstein (1995) argues that QR can be subsumed under the feature-checking operations of Minimalism. A separate QR operation thus becomes unnecessary. This is clearly an elegant solution to the problem of scope determination. The account also extends to cover the scope phenomena found in Chinese and Japanese QP-QP sentences, although as with Aoun & Li's (1993) account, the controversial assumption must be made that subjects in Chinese and Japanese lack subject-raising. However, it appears that the account does not readily extend to cover the lack of pair-list readings in scrambled Japanese [*wh*-object...QP-subject...] questions.

From the point of view of L2 acquisition of scope phenomena in Japanese, the key problem for English learners of Japanese acquiring the scope rigidity of Japanese SOV QP-QP sentences is basically the same as in Aoun & Li's (1993) account, namely the resetting of a VP-internal subject parameter. Specifically, under Hornstein (1995), learners must come to know that Japanese subjects originate in AgrSP, unlike English subjects, which raise to AgrSP from VP. As already explained with reference to Aoun & Li (1993), there is likely to be little empirical evidence in the Japanese input to motivate the relevant parameter resetting. Thus, under this account, too, English learners of Japanese may fail to overcome the poverty-of-the-stimulus problem represented by acquisition of the lack of object-wide scope in Japanese SOV [QP… QP…] sentences. The account does not account for the lack of pair-list interpretations in Japanese [*wh*-object...QP-subject...] questions, therefore no prediction can be made about whether English-speaking and Chinese-speaking learners of Japanese can acquire this lack.

A number of problems for Hornstein's (1995) analysis are noted by Reinhart (1995, 1997). Most importantly, the analysis of QP-QP sentences does not account for scope ambiguity when the subject is a universal QP, as in (72).

72. Every man kissed a woman.

This sentence is true when every man kissed a different woman (subject-wide reading) and when there is one woman whom every man kissed (object-wide reading). However, if definite NPs cannot be interpreted in VP, as claimed above for (61), this means that there can be no LF representation in which the object c-commands the subject (i.e., object in AgrOP, subject in VP), and therefore no object-wide scope. Hornstein (1995: 237–8 fn. 12) argues that this is in fact the case: (72) has only a subject-wide reading, but this reading is indeterminate with regard to how many women each man kissed. He notes that there are multiple readings of (72): if there are 100 men and 50 women, as long as each man kisses one woman, the sentence is true regardless of how many of the women remain unkissed. Thus he suggests that these multiple readings—which all entail the subject-wide reading—are in fact variations of the subject-wide reading. Reinhart (1995, cited in Szabolcsi 2001: 626–7) points out, however, that this reasoning does not cover the ambiguity of examples such as (73).

73. Most but not all of the students attended some seminar.

Here, the subject NP is definite, so in the LF representation, the copy within VP must delete. The remaining structure will allow only the subject-wide reading. This is problematic, since an object-wide reading obtains, in which there is one particular seminar such that most of the students attended it. This reading is not entailed by the subject-wide reading: 'most of the students attended a seminar, but they didn't necessarily attend the same one'. The object-wide reading thus cannot be ascribed to vagueness.

A different problem with Hornstein (1995) is one that affects all of the accounts examined so far: the fact that they deal only with the quantifiers *every* and *some*, or equivalents of those quantifiers in Chinese and Japanese. In fact, there are numerous other quantifiers, as the references above to *most*, *all* and numeral quantifiers has touched upon. The scopal properties of the different quantifiers often do not exactly match those of *every* and *some*. The next section focuses on analyses of scope interpretation that attempt to explain the different scope-taking properties of different quantifiers.

2.3. 'All quantifiers are not equal': analyses of the different properties of different quantifiers

McCawley (1977: 372) observes that '[t]extbooks of formal logic can easily give the casual reader the impression that there are only two quantifiers: "the" universal quantifier [i.e., *every*] and "the" existential quantifier [i.e., *some*]' (brackets and their contents added by the present author). The same comment could be applied to the syntactic accounts of quantification detailed so far in this chapter. However, as

McCawley (1977) and others (e.g., Lakoff 1972; Vendler 1967) have shown, different quantifiers produce different semantic interpretations. Vendler, for example, notes that *all* refers collectively to a group of individuals, while *each* or *every* refers distributively to a group of individuals (Vendler 1967: 74). This is exemplified as follows (Vendler 1967: 72–3):

74. All those blocks are similar.
75. Each (every one) of those blocks is similar.

The sentence in (74) is logically complete, with the meaning that every block in the set is similar to every other block in the set. However, the sentence in (75) lacks a complement: it leaves open the question, 'Similar to what?' (Vendler 1967: 73). Even when a complement is added, the meaning is not the same is in (74):

76. Each of those blocks is similar to every other.

The sentence in (76) is true when each block shares some feature in common with every other block, without entailing that all the blocks are similar.[47] In other words, (74) entails (76), but (76) does not entail (74) (Vendler 1967: 73). The notion of similarity applies to the whole set collectively with *all* in (74), but to subsets of the whole, with *each* or *every* in (76). This is a function of the specific quantifiers.

Since the collective use of *all* entails the distributive use, it is not surprising that examples of ambiguity can be found. The following are from Lakoff (1972: 554 (2), (3)):

77. All the boys carried the couch upstairs.

78. Every boy carried the couch upstairs.

---

[47] Vendler exemplifies this as follows, citing the idea of Goodman (1951) of an 'imperfect community':

> Take three elements with characteristics distributed as follows: *ab*, *bc*, *ac*. Then, with the given interpretation, any two elements will be similar to each other without all of them being similar, since there is no common characteristic running through the total population.
>
> (Vendler 1967: 73)

The sentence in (77) can mean that the boys collectively carried the couch upstairs, or it can mean that each boy individually carried the couch upstairs. (78), on the other hand, has only the latter meaning.

The differences between different types of quantifier affect the interaction of quantifiers in doubly-quantified sentences. Beghelli & Stowell (1997) note that only *each* and *every* can quantify over a singular indefinite QP with the distributive reading (example based on Beghelli & Stowell 1997: 90 (20)):[48]

79. a. Every student read a different book.
    b. Each (of the) student(s) read a different book.
    c. *All the students read a different book.
    d. *Five students read a different book.

Compare (79c) and (79d) with the following, where the object is a plural indefinite:

80. a. All the students read different books.
    b. Five students read different books.

In (80a–b), a distributive reading (e.g., Student A read Book A, Student B read Book B, etc.,...) is the only reading available.[49] The contrast between (79c–d) and (80a–b) shows that *all the* and numbers can be distributive quantifiers when the distributee is a plural noun, but not when it is a singular noun.

A further example of different quantifier types affecting the interactions of quantifiers occurs in sentences such as (81a–b) ((81b is from Reinhart 1978: 124 (36a)).

81. a. Some tourists visited every museum.
    b. Some tourists visited all the museums.

Ioup (1975) observed that the inherent ability of quantifiers to take inverse scope appears to be on a scale, with *every* taking inverse scope more readily than *all*. This seems indeed to be true for (81a–b): the object-wide scope interpretation, whereby

---

[48] Note that (79c) and (79d) are grammatical with collective interpretations, i.e., if 'a different book' is interpreted as being the same book for each student, as in (i):

i. The literature teacher recommended Mary Shelley's *Frankenstein*, but all the students/five students read a different book, namely Bram Stoker's *Dracula*.

[49] Note that the variation whereby each student may have read more than one book (i.e., Student A read Books A, B & C, Student B read Books A & D, etc. ...) is also a distributive reading.

for each museum some tourists visited it, is readily available for (81a), in which the object quantifier is *every*, but is much harder to derive in (81b), in which the object quantifier is *all*. This observation receives further empirical confirmation by Lee, Yip & Wang (1999b) (see Chapter 3) and also by research presented in Chapters 4 and 5 of the present dissertation. In other words, the examples in (79)–(81) show that the distributive function of *all* is not always readily accessible, in contrast to *every*.

Like English, Japanese also has a number of different universal quantifiers in addition to *daremo* 'everyone', which was used in most of the Japanese examples in the previous section. A (non-comprehensive) list is presented in (82), with the translations given in *Kenkyusha's New College Japanese-English Dictionary, 5th Edition* (Collick, Dutcher, Tanabe & Kaneko (eds.) 2002):

82. a. *daremo* 'everyone'
  b. *dono N-mo* 'every N', 'all the N'
  c. *subete(-no N)* 'all', 'everything'; 'all (the N)', 'every (N)'
  d. *minna* 'all', 'everything', 'everyone'
  e. *sorezore(-no N)* 'each'; 'each (N)'
  f. *kaku N* 'each N', 'every N'

The QPs in (82a–b) differ from (82c–f) in that they belong to a distinct class of quantifier occurring cross-linguistically, in which quantificational force derives from a *wh*-pronoun in combination with a quantificational particle (henceforth *wh*+QPt quantifiers, where 'QPt' means 'quantificational particle').[50] The Japanese existential *dareka* 'someone' is also of this type. The morphological analyses of (82a–b) and of *dareka* are shown in (83).[51]

83.

	QP	*wh*-pronoun	+	quantification particle	meaning
a.	*daremo:*	*dare* 'who'	+	*mo*	'everyone'
b.	*dono N-mo:*	*dono* 'which'	+N+	*mo*	'every N'
c.	*dareka:*	*dare* 'who'	+	*ka*	'someone'

[50] Other languages exhibiting *wh*+QPt quantifiers include Croatian, Polish, Malayalam, and, of the other languages investigated in this dissertation, Korean.

[51] Japanese also has other *wh*+QPt quantifiers, including *nanimo* 'everything' (*nani* 'what' + *mo*), *nanika* 'something', and *dokoka* (*doko* 'where' + *ka*). The particle *mo* is a conjunctive particle (e.g., *Hanako-mo Taro-mo* 'both Hanako and Taro'); and *ka*, a disjunctive particle (e.g., *Hanako-ka Taro(-wa)* 'Hanaka or Taro(-TOP)'). For detailed analysis of Japanese *wh*+QPt quantifiers, see Kuroda (1965) and Nishigauchi (1990); also (among others) Hasegawa (1991, 1993); Kawashima (1994a, 1994b); Nishigauchi (1999).

One semantic difference between universal *wh*+QPt quantifiers and the other universals in (82), is that the former behave as negative polarity items in negative contexts, while the latter do not (see references in footnote 51). Another characteristic of Japanese *daremo* 'everyone' and *dono N-mo* 'every N' is that they do not readily yield a collective interpretation. This was mentioned with respect to *daremo* 'everyone' in Section 2.2.4. Example (84), below, provides an illustration with *dono N-mo* 'every N' (Kawashima 1994a: 136 (44)):[52]

84. Dono gakusei-mo kooen de atta.
every student-QPt park in met
'Every/each student met in the park. (i.e., there were sub-meetings which each student engaged in at the park)'/ *'The students all met in the park.'

Kawashima (1994a) does not compare *wh*+QPt quantifiers with other non-*wh*+QPt quantifiers (such as those in (83c–f)). However, an informal survey by the present author of 10 native-Japanese speakers reveals that, if *dono gakusei-mo* 'every student' in (84) is replaced by *subete-no gakusei(-wa)* 'every student/all the students(-TOP)', the collective interpretation (i.e., 'The students all met in the park') becomes readily available, while the distributive interpretation is not easy to get.[53] Based on this semantic difference, *subete-no N* is henceforth translated as 'all the N' while *dono N-mo* is translated as 'every N'. This is intended to indicate that *subete-no N* is associated with collective interpretations while *dono N-mo* is associated with distributive interpretations; however, it is not intended to indicate direct semantic equivalence between *subete* and *all*, and *dono…mo* and *every*.

A further difference between *subete-no N* and *wh*+QPt universals is that the former can c-command a *wh*-word without making the sentence unacceptable (Tomioka 2004). Thus, *subete-no N* is acceptable as the subject of a (non-scrambled) [QP-subject…*wh*-object…] question (85), whereas *wh*+QPt universals are not (86) (as detailed in Section 2.1 for *daremo* 'everyone'):

---

[52] *Dono N-mo* is glossed as 'every N-QPt' throughout this dissertation.

[53] Some of the informants in this survey commented that they found the collective interpretation applicable also to the sentence in (84), contra Kawashima's claim. This is a concrete example of individual variation with respect to judgements about quantified sentences. In this case, it may arise because, as noted at the beginning of this section, a collective interpretation entails a distributive interpretation. Therefore, it may be conceptually difficult to separate the two meanings. Despite this disagreement by some informants about whether or not the collective interpretation applies to (84), all informants agreed that, with *subete-no gakusei* as the subject, the collective interpretation applies.

85. Subete-no gakusei-ga nani-o katta no?
all-GEN student-NOM what-ACC bought Q
'What did all the students buy?'

86. ??Dono gakusei-mo nani-o katta no?
every student-QPt what-ACC bought Q

Despite the contrast between *subete-no N* and *dono N-mo* shown in (85) and (86), *subete-no N* is like *wh*+QPt universals in that it does not support a pair-list interpretation of questions with a QP-subject and *wh*-object, whether in the SOV order or the scrambled OSV order. This is illustrated for a scrambled [*wh*-object ... QP-subject ...] question in (87) (based on Yusa 1995: 341 (55)):[54]

87. Dono hon-o subete-no gakusei-ga yomimasita ka.
which book-ACC all-GEN student-NOM read Q
'Which book did every student read?'

Answer: *John-wa *LGB*-o, Bill-wa *SS*-o...
John-TOP *LGB*-ACC Bill-TOP *SS*-ACC
'John read *LGB*, Bill read *SS*...'

Yusa (1995) notes that replacing *subete* in (87) with *sorezore* 'each' renders the pair-list reading acceptable.

To summarise, it is clear that, cross-linguistically, specific quantifiers have specific semantic functions that can affect quantifier scope assignment. Uniform analyses of quantifier scope assignment do not capture these differences. The analyses to be presented in the following sections attempt to address the different semantic functions of different quantifiers.

2.3.1. A 'Target Landing Sites' theory of scope

2.3.1.1. English QP-QP sentences in the Target Landing Sites theory

Recall that, in Minimalism, a problem with QR, as proposed by May (1977, 1985), is that it is not feature-driven. Hornstein (1995) solved the problem by getting rid of QR. By contrast, the 'Target Landing Sites' theory of scope by Beghelli and Beghelli & Stowell (Beghelli 1993, 1995, 1997; Beghelli & Stowell 1994, 1997) retains QR, but argues that it is a feature-driven operation.[55] Just as *wh*-phrases have long been

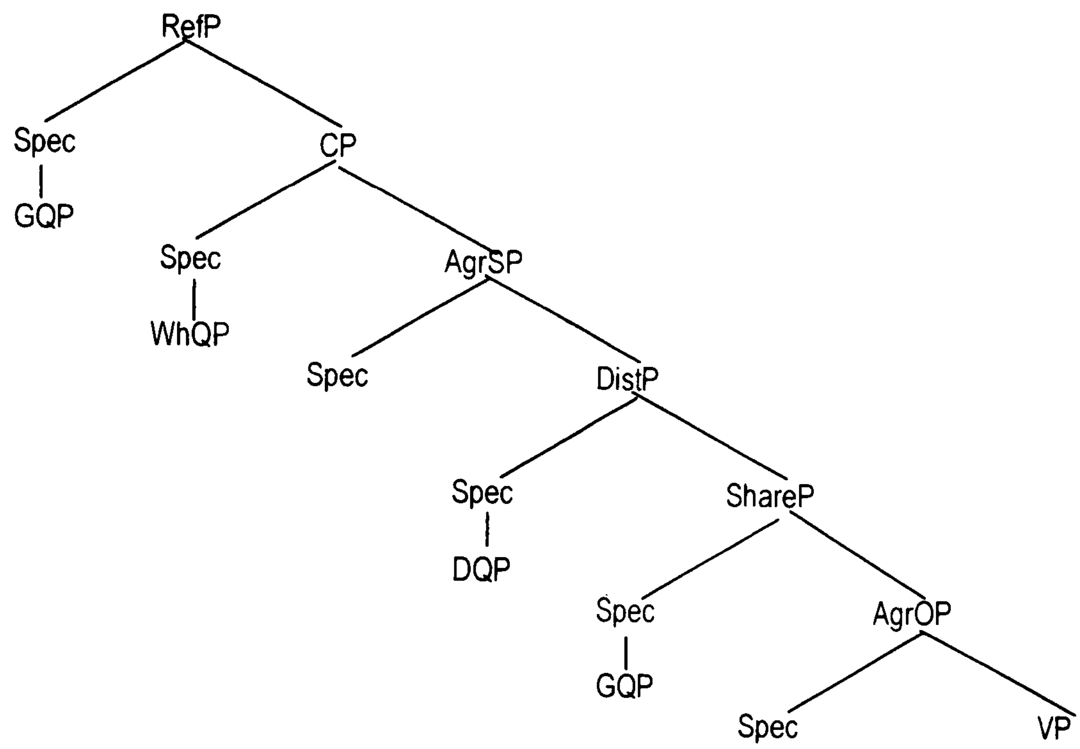
[54] *LGB* = *Lectures on Government and Binding*, *SS* = *Syntactic Structures*.

[55] Beghelli (1995, 1997) refers to the account as a 'Target Landing Sites theory'. Beghelli & Stowell (1997) refer to the same account as 'a checking theory of scope'.

argued to bear a specific feature [+wh] which is checked in a dedicated functional projection, CP, the 'Target Landing Sites' theory proposes that quantifiers also have inherent features which are checked in specific functional projections The theory proposes a classification of QP types, into (among others) 'Distributive-Universal QPs' (DQP) for strongly distributive quantifiers, like *every* and *each*, and 'Group-denoting QPs' (GQP) for, for example, bare numerals, and *some*.[56] DQPs are morphologically singular ([+SG]), and have a [+Distributive] feature which is checked via Spec-head agreement with a distributive operator located in a functional projection, Dist(ributive)P;[57] and GQPs have a [+Group Ref(erent)] feature which is checked with an existential operator occurring in two further projections, Ref(erential)P and ShareP. RefP is argued to be a topic-like position for wide scope interpretations, and ShareP, the position that allows the GQP to serve as the distributed share (distributee) of a DQP (Beghelli 1997: 371). The availability of the different functional projections enables the theory to address the different scope-taking properties of different quantifiers. The clause structure incorporating the proposed functional projections is shown in (88) (Beghelli & Stowell 1997: 76 (2)):[58]

[56] Beghelli (1995, 1997) and Beghelli & Stowell (1997) note that this classification of QP types is largely based on Szabolcsi (1994, 1997).
[57] Beghelli & Stowell (1997) modify this proposal in Section 5 of their paper, arguing that *every* is in fact underspecified for [Distributive]. It has access to Spec,DistP by virtue of its [+SG] feature—a pre-condition for the application of the distributive operator in DistP— and its movement to Spec,DistP is motivated by the binding of its set variable (see Beghelli 1997 for detail). This modification accounts for certain differences in behaviour between *each* and *every*. For the purpose of the present section, however, it is enough to assume the simplification given in the main text: that *every* moves to Spec,DistP for checking purposes. See further discussion in Section 2.3.1.3.
[58] NegP is proposed to occur between ShareP and AgrOP. Since the present dissertation does not address negative quantifiers, NegP is omitted here.

88.



The scope ambiguity of an English sentence such as (89a) is captured by the LF representations in (89b) and (89c).[59] (Curly brackets indicate reconstruction, discussed following the example.)

89. a. Two students read every book.
    b. [RefP Two students$_i$ [AgrSP t_i [DistP every book$_j$ [AgrOP t_j [VP read t_j]]]]]
    c. [AgrSP t_i [DistP every book$_j$ [ShareP {two students}$_i$ [AgrOP t_j [VP read t_j]]]]]

(89b) and (89c) are derived as follows. As in Hornstein's (1995) account, nominative and accusative Case features are checked in Spec,AgrSP and Spec,AgrOP, respectively. The DQP *every book* must also check its [+Distributive] feature, and thus ends up in Spec,DistP. The GQP *two students* must check its [+Group Ref] feature either in Spec,RefP or Spec,ShareP. In the wide-scope interpretation of *two students* ('there are two students such that those two students read every book') represented in (89b), the relevant checking takes place within RefP. In the narrow-scope interpretation ('for every book, there are two students who read it'), in which *two students* provides the distributed share to *every book*, the GQP reconstructs from

[59] I follow Beghelli (1995) and Beghelli & Stowell (1997) in not indicating a VP-internal position for the subject. However, this is an abbreviation: their theory assumes that subjects originate within VP, as indicated in, for example, the fully articulated tree structures in Beghelli (1995: 74 (61)).

Spec,AgrSP to Spec,ShareP. This is represented in (89c). Reconstruction, in this theory, is constrained by the following principle (Beghelli 1995: 78 (64)):

90. *Availability of reconstruction*
A QP may only lower/reconstruct to a scope position where its semantic and morphological features are checked.

In other words, reconstruction is available when the position from which it takes place is not one in which semantic features are interpreted. Thus, reconstruction from Spec,AgrSP is permitted.

The problem of the lack of distributive interpretation of *all* when it occurs in object position is accounted for in this system by proposing that *all* is not a DQP, but a GQP. GQPs are argued to behave as 'pseudo-distributors' given the right syntactic configuration, such as that seen (80), repeated below in (91).

91. a. All the students read different books.
b. Five students read different books.

The distributive function of the GQPs *all* and *five* in (91a–b) is argued (Beghelli 1995, 1997) to arise through a covert distributive operator, in the form of a silent adverbial *each*, assumed to occur in the same AgrSP, AgrOP, or ShareP (below the specifier position) in which the distributor QP (or its trace) is located.[60] RefP is ruled out as a possible site for silent *each* on the basis of the distribution of overt *each*, which cannot occur in a position higher than the canonical subject position:

92. a. *Each the girls ate an apple.
b. The girls each ate an apple.
c. The girls ate an apple each.

The distributee QP must be c-commanded by silent *each*. The resulting LF structures for (91a–b) are presented in (93a–b), respectively (based on Beghelli 1997: 379 (53)). (Irrelevant projections are omitted. Italics are used to indicate that *each* is covert.).

93. a. [$_{\text{RefP}}$ All the students$_i$ [$_{\text{AgrSP}}$ t_i [*each* [$_{\text{ShareP}}$ different books$_j$ [$_{\text{VP}}$ read t_j]]]]]
b. [$_{\text{RefP}}$ Five students$_i$ [$_{\text{AgrSP}}$ t_i [*each* [$_{\text{ShareP}}$ different books$_j$ [$_{\text{VP}}$ read t_j]]]]]

[60] Beghelli (1995, 1997) also lists as AgrI[ndirect]O[bject]P as a possible adjunction site for silent *each*.

In both (93a) and (93b), the conditions of pseudo-distributivity are satisfied: both the subject and object are GQPs, and the distributive operator, silent *each*, occurs in a position (AgrSP) that c-commands the distributee, *different books*.

The lack of inverse scope interpretation when a GQP occurs in object position falls out from the assumption that covert *each* cannot occur above Spec,AgrSP. Consider again (81b), repeated below as (94).

94. Some tourists visited all the museums.

In the inverse scope interpretation of (94), 'for each museum, some tourists visited it', *all the museums* acts as distributor and must take scope in RefP, above the subject *some tourists*. However, since covert *each* cannot occur in RefP, there is no way for it to c-command *some tourists*. Thus the distributee fails to be c-commanded by the distributive operator, and the lack of object-wide scope of *all the museums* is captured.

2.3.1.2. Japanese QP-QP sentences in the Target Landing Sites theory

Beghelli (1995, 1997) and Beghelli & Stowell (1997) do not address the scope rigidity of Japanese and other languages. If Japanese *daremo* 'everyone' and *dono N-mo* 'every N' are assumed to be DQPs, the lack of inverse scope when these expressions are in object position is not predicted. Appealing to the VP-internal subject hypothesis, as in Aoun & Li (1993) and Hornstein (1995), does not solve the problem, since in the Target Landing Sites account, scope assignment is a result of the movement of quantifiers to dedicated positions in order to check scope-related features. Thus a distributive universal quantifier in Japanese has to move to Spec,DistP, and a group-denoting quantifier has to move to Spec,RefP or Spec,ShareP regardless of whether or not they originate in VP. Therefore the sentence in (95a) is predicted to be ambiguous in the same way as its English counterpart in (89). (95b) shows the LF representation for an inverse scope reading of (95a).[61]

95. a. Hutari-no gakusei-ga dono hon-mo yonda. S>O, *O>S
two.CL-GEN student-NOM every book-QPt read
'Two students read every book.'

[61] Quantitative data presented in Chapters 4 and 5 confirm that Japanese sentences like (95), with a numerically quantified subject and universally quantified object, tend to lack an object-wide reading.

b. $[_{AgrSP}\ t_i\ [_{DistP}$ dono hon-mo$_j$ $[_{ShareP}$ {hutari-no gakusei-ga}$_i$ $[_{AgrOP}\ t_j\ [_{VP}\ t_j$ yonda]]]]]

In (95b), the object *dono hon-mo* 'every book' in DistP c-commands the subject *hutari-no gakusei* 'two students' in ShareP, meaning that an object-wide scope reading is predicted to be available.

One way of ruling out the inverse scope representation for (95) would be to propose that *dono N-mo* 'every N' is not a distributive universal, but is instead a pseudo-distributor, like English *all.* A diagnostic for DQPs is that they must be morphologically singular. In fact, it is not at all clear whether *dono N-mo* meets that requirement. There is generally no singular/plural distinction for nouns in Japanese.[62] However, an optional plural marker *tachi* can be affixed to human nouns, as in (96):[63]

96. Kodomo-tachi-wa kooen de asonda.
child-PLURAL.HUMAN-TOP park in played
'The children played in the park.'

If this plural marker is used with *dono N-mo*, as in (97), native Japanese opinions are divided about the acceptability.

97. Dono kodomo-tachi-mo kooen de asonda.
every child-PLURAL.HUMAN-QPt park in played
*'Every children played in the park.'

Some native Japanese speakers find (97) acceptable, while others claim it is ungrammatical precisely on the grounds that *dono N-mo* and *tachi* are incompatible.[64] Note that, by contrast, *tachi* is unanimously judged perfectly acceptable with another Japanese universal quantifier, *subete-no N* 'all the N':

98. Subete-no kodomo-tachi-ga koen de asonda.
all-GEN child-PLURAL.HUMAN-NOM park in played
'All the children played in the park.'

---

[62] Korean and Chinese also do not make a singular/plural distinction.
[63] If *tachi* is omitted in (96), the sentence is ambiguous with regard to whether there is one child or more than one.
[64] The judgements about the data in (97), (98) and (100) come from at least five native Japanese linguists.

Another test for the singular distributive properties of a universal quantifier is whether or not it can be the subject of a verb requiring a semantically plural agent such as *surround*. Beghelli & Stowell (1997: 88) illustrate that *every* is not readily acceptable as the subject of *surround*, in contrast to *all* (Beghelli & Stowell 1997: 88 (17a–b)):

99. a. ?Every boy surrounded the fort.
    b. All the boys surrounded the fort.

They explain that the distributive *every* forces a construal whereby the 'surrounding event' is attributed individually to each member of the subject set (i.e., each boy in (99))—a reading that is incompatible with the semantics of *surround*, which requires a collective construal of the subject. Considering the Japanese (approximate) equivalents of (99), given in (100), native Japanese opinions are again divided about the acceptablity of the example with *dono N-mo* 'every N' (100a), while the example with *subete-no N* 'all the N' (100b) is unanimously judged acceptable.

100. a. Konsaato-no ato, dono fan-mo sutaa-o torimaita.
     concert-GEN after every fan-QPt star-ACC surrounded
     ?'After the concert, every fan surrounded the star.'

  b. Konsaato-no ato, subete-no fan-ga sutaa-o torimaita.
     concert-GEN after all-GEN fan-NOM star-ACC surrounded
     After the concert, all the fans surrounded the star.'

Some speakers reject (100a) while others find it perfectly acceptable.[65]

The lack of consensus about (97) and (100a) is perhaps to be expected in a language that does not make a singular/plural distinction. Clearly, *dono N-mo* is not obviously singular, unlike English *every*. Let us assume, then, that *dono N-mo* does not have the feature [+SG] and is therefore excluded from Spec.DistP. As such, *dono N-mo* can be classed as a GQP rather than a DQP, and hence, can act as a pseudo-distributor. Since pseudo-distributors cannot take wide scope in object position, Japanese scope rigidity in canonical QP-QP sentences is accounted for.

[65] In Marsden (2003: 499 fn. 4) I claimed that the sentence in (100a) was unacceptable in Japanese, a claim that was based on the judgements of three native Japanese speakers. Since then, I have surveyed more native speakers of Japanese and therefore revise the claim, as described here.

The ambiguity of scrambled sentences such as (101) can also be accounted for if it is assumed that the covert distributive operator *each* can scramble with the object to the landing site of scrambling above AgrSP.

101. Dono hon-mo hutari-no gakusei-ga yonda. S>O; O>S
every book-QPt two.CL-GEN student-NOM read
'Two students read every book. (*scrambled*)'

This is a reasonable assumption, given that Japanese floating quantifiers can occur overtly above the subject when the object with which they are associated is scrambled. This is illustrated with *hitori-hitori* 'each' in (102b), a scrambled version of (102a):[66]

102. a. Sensei-ga kodomo-o hitori-hitori sikatta.
teacher-NOM child-ACC each scolded
'The teacher scolded each child.'

b. [[Kodomo-o]$_i$ hitori-hitori [AgrSP sensei-ga t_i sikatta]]

For (101), the distributive object-wide reading is accounted for by the following LF representation. (The landing site of scrambling is designated as XP.)

103. [RefP dono hon-mo$_j$ [XP t_j [*each* [AgrSP t_i [ShareP {hutari-no gakusei-ga}$_i$ [AgrOP t_j [VP t_j yonda]]]]]]]

In (103), *each* c-commands *hutari-no gakusei* 'two students' from within XP. Pseudo-distributivity can thus apply, with *dono hon-mo* 'every book' distributing over *hutari-no gakusei*.

To summarise, the Target Landing Sites theory accounts for the facts of Japanese QP-QP scope interpretation, provided that the following empirically motivated assumptions are made: (i) Japanese universal quantifiers are not DQPs, due to their apparent lack of a [+SG] feature; instead they are GQPs and as such pseudo-distributors; (ii) the covert distributive operator can occur above AgrSP if the QP with which it is associated has scrambled above the subject.

[66] In (102b), I leave open the question of whether the floated quantifier scrambles with the object from a VP-internal position, or whether it is base-generated above AgrSP, since this is irrelevant to the argument.

2.3.1.3. *Wh*-QP questions in the Target Landing Sites theory

Beghelli (1997) explores the question of pair-list readings in detail, with particular reference to differences between *each* and *every*. This section will present a brief summary of his account as it relates to the availability of pair-list readings in [*wh*... *every*...] questions. As already seen, a pair-list reading is possible in English when *every* is the subject of a *wh*-question, but not when it is the object (Beghelli 1997: 349, (1) & (2)):

104. What did every student read?
 'for every student *x*, what did *x* read?'

105. Who read every book?
 *'for every book *x*, who read *x*?'

Beghelli notes that this subject-object asymmetry parallels the subject-object asymmetry of pseudo-distributivity in declaratives: a pseudo-distributor like *all* can take wide-scope when it is in subject position in a declarative sentence, but not when it is object position. Beghelli accounts for pair-list readings with *every* by suggesting that, in the pair-list interpretation, *every* behaves as a pseudo-distributor, not a strong distributor (since, if it were a strong distributor in (105), the *every*>*who* reading should be available, contra fact).

The ability of *every* to behave as a pseudo-distributor comes from a modification to the theory of *every* (already mentioned in footnote 57): *every* is underspecified for [Distributive], instead of bearing a [+Distributive] feature. Consequently, it is not forced to move to DistP to check a [+Distributive] feature. Its movement to DistP is a consequence of an assumed feature of distributive quantifiers (following Szabolcsi 1997): that a distributive quantifier introduces a discourse referent in the form of a set variable.[67] This variable must be bound, and in a declarative sentence it is bound by the existential operator in RefP. Binding by the existential operator is assumed to drive *every* to move to Spec,DistP (Beghelli 1997: 383). In questions, however, the question operator in CP is closer to *every* than the existential operator in RefP. Beghelli argues that this opens the possibility that the set variable of distributive *every* be bound by the question operator. When the set variable of *every* is bound by the question operator, *every* remains in its Case

[67] For example (Beghelli 1997: 371), 'the QP *every man* introduces a variable X to be assigned the set containing all the men in the situation'.

position, and does not move to Spec,DistP. The pair-list reading emerges when the *wh*-phrase can reconstruct to a position under the scope of *every*. The representation for (104) is given in (106) (based on Beghelli 1997: 393, (74); 'Q-Op' = 'question operator'):

106. [CP [Q-Op$_i$ + t_i] [AgrSP [every student]$_{k/i}$ [AgrOP {what}$_i$ [VP read t_i]]]]

In (106), *what* (but not the question operator) has reconstructed from CP to its Case position, AgrOP, where it is bound by the question operator. *Every student* intervenes in the binding path between the question operator and *what*, and thus its set variable is argued to also be bound by the question operator. In this configuration, pseudo-distributivity applies: *every* can act as a distributor because its set variable is bound and it c-commands and thus can take scope over the distributee, *what*, yielding the pair-list interpretation. However, when the *wh*-phrase is in subject position and *every N* in object position, as in (105), no pair-list representation is available. This is because there is no position under the scope of *every* in AgrOP to which *who* could reconstruct.[68] The LF representation is given in (107) (based on Beghelli 1997: 391 (71i)):

107. [CP [Q-Op$_i$ + who$_i$] [AgrSP t_i [AgrOP [every book]$_k$ [VP read t_k]]]]

The conditions for pair-list interpretation are not fulfilled in (107): the set variable of *every N* is not bound, so *every N* cannot act as a distributor; in addition, *who* cannot move to a position under the scope of *every*.

Thus, the Target Landing Sites theory accounts for the lack of pair-list readings of English [*wh*-subject...QP-object...] questions and the availability of pair-list readings in English [*wh*-object...QP-subject...] questions. However, as with Hornstein's (1995) account, the lack of pair-list readings in scrambled Japanese [*wh*-object... QP-subject...] questions is problematic. There seems to be no reason why a Japanese *wh*-object, such as *nani* 'what' in (108a), should not be able to reconstruct to its Case position in AgrOP, yielding the LF representation given in (108b):[69]

---

[68] The VP-internal subject position is ruled out as a possible reconstruction site by the constraint on the availability of reconstruction given in (90), since Spec,VP is not a position where semantic or morphological features are checked.

[69] *Wh*-movement of *nani* 'what' to CP is assumed to occur covertly, since Japanese is a *wh*-in-situ language.

108. a. Nani-o$_i$ daremo-ga t_i katta no. *wh*>QP; *QP>*wh*
what-ACC everyone-NOM bought Q
'What did everyone buy?'

b. [CP [Q-Op$_i$ + t_i] [XP t_i [AgrSP [daremo]$_{k/i}$ [AgrOP {nani-o}$_i$ [VP t_i katta]]]] no]

In the structure in (108b), the conditions for a pair-list interpretation appear to be met: *nani* 'what' is under the scope of *daremo* 'everyone', and *daremo* 'everyone' is bound by the question operator and therefore should be able to act as a distributor. Further development of this theory is thus required in order to account for the lack of pair-list readings in scrambled Japanese [*wh*-object ... QP-subject ...] questions.

2.3.1.4. Comment

The Target Landing Sites theory of scope shares with Hornstein's (1995) uniform theory of quantifier scope that scope assignment is the result of feature checking. It differs from previous analyses in that it offers an account for the different scope-taking properties of different types of quantifier, by proposing specific functional projections for specific quantifier types. It appears that the theory can be applied successfully to the facts of Japanese QP-QP interpretation, if it is assumed that universal quantifiers in Japanese do not fall into the category of DQPs like English *every*. This assumption is supported by empirical evidence showing that *dono N-mo* 'every N' is not readily construed as singular—a pre-condition for classification as a DQP. However, the theory does not readily account for the lack of pair-list readings in scrambled Japanese [*wh*-object ... QP-subject ...] questions. Consequently, only QP-QP sentences are considered in the following discussion of the L2 acquisition of Japanese scope interpretation under the Target Landing Sites theory.

For English-speaking learners of Japanese, the key difference between English and Japanese, under this theory, is that Japanese universal quantifiers (at least those investigated: *dono N-mo* and *subete-no N*) are not inherently singular whereas English *every* is. This presumably comes from the fact that Japanese, unlike English, does not make a syntactic distinction between singular and plural. The acquisition problem is therefore to internalise this lack of singular/plural distinction so that it becomes part of the interlanguage grammar and applies to all NPs. Assuming that the categories of DQP, GQP, etc. are part of UG, knowledge that Japanese NPs cannot be singular will automatically rule out Japanese universal quantifiers from being classed as DQPs and being able to take inverse scope in a canonical SOV doubly-quantified sentence.

Evidence for the lack of singular/plural distinction in Japanese is likely to be available in the input in the form of nouns occurring in both singular and plural contexts without any morphological indication of number. In other words, in the English-based interlanguage, plural morphology is 'expected' in a plural NP. No plural morphology occurs in this required context in the Japanese input. This lack of the expected morphology in a required context constitutes indirect 'negative evidence', which may trigger acquisition of the lack of singular/plural distinction in Japanese.[70, 71] Given such evidence, English-speaking learners are predicted to be able to acquire the lack of object-wide scope in Japanese SOV QP-QP sentences as a corollary of acquiring the knowledge that Japanese NPs are neither singular nor plural. However, at least two factors may impede the clarity of the indirect negative evidence for the lack of singular/plural distinction in Japanese. First, the L1 English grammar can parse nouns without singular or plural morphology: for instance, a sentence such as *I like cake* (v. *I like cakes*) is grammatical if *cake* is construed as a mass noun. Therefore, NPs without singular or plural morphology could conceivably be represented by an English-based interlanguage grammar without immediately motivating restructuring of that grammar to one in which nouns are underspecified for number. Second, the fact that Japanese sometimes makes use of a plural suffix on human nouns (as described above) further blurs the evidence. Therefore, there is likely be at least some delay in the restructuring of the English-based interlanguage

---

[70] 'Negative evidence' is defined as information about the impossibility of an utterance (Schwartz & Gubala-Ryzak 1992). Note that the potential exploitation of indirect negative evidence to acquire the lack of singular/plural distinction in Japanese does not contradict the claim that the unambiguity of Japanese SOV QP-QP sentences cannot be acquired by 'noticing' that Japanese lacks object-wide scope. This is because there is no clearly definable, finite required context for object-wide scope in a QP-QP sentence—unlike for singular or plural morphology in an NP. For example, either of the sentences in (i) and (ii) could be used in English to express 'for each book, someone read it'.

i. Someone read every book.
ii. Every book was read by someone.

There is no 'required context', in normal discourse, in which (i) must be used instead of (ii). Consequently, there is no way for English learners of Japanese to 'notice' that a Japanese equivalent of (i) (i.e., an active, transitive, SOV QP-QP sentence) is not used to convey the meaning 'for each book, someone read it'.

[71] In addition, adult, classroom-based learners of Japanese are likely to receive instruction, early on in their course, about the lack of singular/plural distinction. I leave open the question of whether such instruction-derived meta-linguistic knowledge could lead to the relevant grammar becoming part of the learner's L2 competence. (See Schwartz & Gubala-Ryzak (1992) for argument that explicit negative evidence via instruction does not lead to restructuring of the interlanguage.)

grammar to one in which there is no singular/plural distinction. Consequently, if learners of an early enough stage are investigated, this account predicts that English-speaking learners will differ from Korean-speaking and Chinese-speaking learners with respect to Japanese scope interpretation: Korean-speaking and Chinese-speaking learners of Japanese will have knowledge of Japanese scope rigidity from the outset as a corollary of L1-derived knowledge that Japanese lacks a singular/plural distinction; English-speaking learners of lower proficiency levels will initially allow non-native-like scope ambiguity as long as the L1-derived singular/plural distinction remains unchanged. However, given that evidence is available to motivate the relevant restructuring of the English-Japanese interlanguage, more advanced English-speaking learners of Japanese are predicted to disallow object-wide scope in Japanese SOV QP-QP sentences.

2.3.2. Extra-syntactic accounts of quantifier scope

All the accounts of quantifier scope presented thus far in this chapter have been syntactic. However, a number of accounts explore the roles of pragmatic, semantic, and other processing principles, in addition to syntax. This section outlines some of these accounts, focusing in particular on analyses that deal specifically with Japanese. Like the Target Landing Sites theory detailed in the previous section, the analyses discussed here attempt to account for the different scopal properties of different quantifiers. In addition, they address the issue of individual differences in judgements about quantifier interpretation.

2.3.2.1. Scope ambiguity resolution via multiple processing principles

The syntactic accounts of scope interpretation discussed in the previous sections all appeal to the principle of c-command in some form: the narrow-scope quantifier is c-commanded by the wide-scope quantifier (or by a quantificational operator) at some level of representation. Other accounts of scope interpretation have identified other factors as contributing towards ambiguity resolution. Ioup (1975), for example, proposed the quantifier hierarchy, given in (109):

109. *each* > *every* > *all* > *most* > *many* > *several* > *some* (+ N [PL]) > *a few*

The rationale behind the hierarchy is that it places distributive quantifiers highest, as most likely to take wide scope, with other quantifiers then arranged in order of the size of set they specify (i.e., *all* specifies a larger set than *most*), since the tendency to take wide scope seems to decrease with decreasing set size.

Another hierarchy that appears relevant to scope ambiguity resolution is the thematic hierarchy (Grimshaw 1991; Jackendoff 1972), whereby agents are more prominent than experiencers, which are more prominent than themes. This is manifested in scope interpretation in Japanese by the fact that the locative object of an unaccusative verb can take (inverse) scope over the experiencer subject, as shown in (110) (Yatsushiro 1996: 324 (8b)):

110. Daremo-ga dokoka-ni tuita. S>O-LOC; O-LOC>S
everyone-NOM somewhere-LOC arrived
'Everyone arrived somewhere.'

The availability of (locative-)object-wide scope in (110) clearly contrasts with the lack of object-wide scope in Japanese SOV QP-QP sentences with transitive verbs. The initial stage of pilot testing conducted for the experimental part of this dissertation provides empirical evidence of this contrast (see Chapter 4). Experimental investigation by Lee, Yip & Wang (1999a, 1999b) (see Chapter 3) shows that the same contrast holds in Chinese: quantified locative objects are readily interpreted as taking inverse scope over quantified experiencer subjects, whereas as quantified theme objects are not readily interpreted as taking inverse scope over quantified agent subjects.

A third factor that has been proposed as playing a role in scope ambiguity resolution is linear order (e.g., Fodor 1982; Kroch 1975). Specifically, quantifiers closer to the left of the surface order of a sentence tend to take scope over quantifiers further to the right. Experimental work by Kurtzman & MacDonald 1993 (see Chapter 3) and by the present author (see Chapters 4 and 5) on ambiguous English QP-QP sentences confirms that the 'left QP > right QP' interpretation appears easier to get than the inverse scope interpretation, despite the fact that both are possible.[72]

[72] The fact that left-most QPs tend to take scope over right-most QPs may be connected to the fact that pragmatic topics tend to occur at the beginning of a sentence. The role of (pragmatic) topic-hood in quantified scope interpretation is discussed in more detail in Section 2.3.2.2, following.

An example of an account of quantifier scope interpretation that makes use of factors such as those described above (quantifier hierarchy, thematic hierarchy, and linear order) is the 'functional' account by Kuno (1991), Kuno & Takami (2002), and Kuno, Takami & Wu (1999, 2000). In fact, this account entails a number of problems—as described, below—which mean that exploration of its implications for the L2 acquisition of Japanese scope phenomena is not pursued. Nonetheless, since substantial sections of the analysis are devoted to scope phenomena in Japanese, a thesis on the L2 acquisition of these phenomena is not complete without at least mentioning the account. A brief outline thus follows.

Kuno *et al.* propose a number of scope-determining processing principles, such as those listed in (111) (based on Kuno 1991: 269–274; Kuno *et al.* 1999: 79–80; Kuno & Takami 2002: 203–204).[73]

111. a. Subject QP > Object QP
     b. Human QP > Non-human QP
     c. Lefthand QP > Righthand QP
     d. i. Quantifier hierarchy for English: *each* > *some* (+ N [SG]) > *every* > *all* > *most* > *many* > *several* > *some* (+ N [PL]) > *a few*
        ii. Quantifier hierarchy for Japanese: 'each' > other quantifiers
     e. (Japanese/Chinese-specific): Universal QP > Existential QP

The *Subject QP > Object QP* principle (111a) and the *Human QP > Non-human QP* principle (111b) can both be seen as relating to the thematic hierarchy: subjects are often agents, and humans are often agents; thus, it is not surprising that subject QPs and human QPs should tend to take wide scope. The *Lefthand QP > Righthand QP* principle (111c) is clearly a linear processing principle. The language-specific quantifier hierarchy principles in (111d) are based on Ioup's (1975) quantifier hierarchy.[74] Finally, the Japanese/Chinese-specific *Universal QP > Existential QP* principle in (111e) appears to be a further extension of the quantifier hierarchy.

The various principles are claimed to interact in a type of parallel-processing system. Specifically, the principles 'cast votes' within an expert system, which determines which quantifier in a QP-QP sentence takes wide scope. For example, the

[73] Kuno *et al.*'s comprehensive list of principles includes additional principles relating to specificity, topic-hood, speaker/hearer v. third person distinction, degree of discourse-linking of the QPs, and other factors. For the purposes of the present discussion, the principles listed in (111) suffice.
[74] Kuno *et al* argue that singular *some* should occur between *each* and *every* in the quantifier hierarchy for English (111c.i). However, Ioup's (1975) quantifier hierarchy omits singular *some* on the grounds that it is exceptional, specifying a set of only one, yet tending to take wide scope.

Subject QP > Object QP principle (111a) leads to assignment of one vote in favour of a wide-scope interpretation to a subject QP, and none to an object QP. In addition to the principles in (111), the expert system gives one 'baseline vote' to each QP, under the assumption that 'given that two quantifiers interact scopally in a sentence, each quantifier has some chances, however remote, of claiming a wide-scope interpretation' (Kuno & Takami 2002: 205). The vote tallies for each quantifier in the sentence in (112) are thus as shown in (113) (based on Kuno & Takami 2002: 220 (71b)):

112. Daremo-ga dareka-o aisite-iru. S>O; *O>S
everyone-NOM someone-ACC loves
'Everyone loves someone.'

113.

daremo 'everyone'	*dareka* 'someone'
Baseline	Baseline
Subject QP	Human QP
Human QP	
Lefthand QP	
Universal QP	
5 votes	2 votes

Verdict: unambiguous (*daremo*-wide scope : *dareka*-wide scope, 5:2)

In (113), *daremo* 'everyone' gets five votes as a possible wide-scope-taking element, while *dareka* 'someone' gets two. This difference in vote tallies is claimed to predict that *daremo* will unambiguously take wide scope over *dareka*. Thus, inverse scope is ruled out, in accordance with the claims (e.g., Hoji 1985) about Japanese scope interpretation in SOV QP-QP sentences.

However, Kuno *et al.*'s account appears to be problematic for other data. When the subject is an existential QP and the object a universal, as in (114), the system is claimed to predict ambiguity, contra the general assertions of scope rigidity for Japanese. The vote tally is shown in (115) (based on Kuno & Takami 2002: 120, (71a)).

114. Dareka-ga daremo-o aisite-iru.
someone-NOM everyone-ACC loves
'Someone loves everyone.'

115.

dareka 'someone'	*daremo* 'everyone'
Baseline	Baseline
Subject QP	Human QP
Human QP	Universal QP
Lefthand QP	
4 votes	3 votes

Verdict: ambiguous (*dareka*-wide scope : *daremo*-wide scope, 4:3)

In Kuno *et al.*'s system, a difference of just one vote, as in (115), predicts ambiguity: either QP may take wide scope, although the preferred interpretation will be wide scope for the QP with the highest tally. This result clearly contradicts the claim of Hoji (1985) and others that Japanese lacks object-wide scope in doubly-quantified SOV sentences. However, Kuno *et al.* assert that this is indeed the correct result, and that—contra Hoji (1985)—a canonical Japanese sentence with an existential QP subject and universal QP object, such as (115), is fundamentally 'ambiguous for most speakers' (Kuno *et al.* 1999: 105). They argue that their system nonetheless allows for the judgements of those individuals who find a sentence like (114) unambiguous. Such individual variation, they claim, could be the result of idiolectal variations in the relative weights of the different principles (Kuno & Takami 2002: 211). This means that the judgement by some Japanese speakers that (114) lacks an inverse scope reading could be accounted for by suggesting that the role of the *Universal QP > Existential QP* principle is diminished in the idiolect of those speakers, or perhaps the importance of the *Lefthand QP > Righthand QP* principle is increased.

Kuno *et al.*'s claims are further complicated by their acknowledgement that, unlike (114), the [∃-NOM...∀-ACC...] sentence in (116) is unambiguous, with no object-wide scope reading. The sentence in (116) is an example frequently used by Hoji (1985). It differs from (114) in that the verb is *semeta* 'criticised', instead of *aisite-iru* 'loves'.

116. Dareka-ga daremo-o semeta. S>O; *O>S
someone-NOM everyone-ACC criticised
'Someone criticised everyone'.

Kuno & Takami (2002: 222) ascribe the lack of ambiguity in (116) to pragmatic implausibility of a distributive reading of *daremo* 'everyone' in this sentence: 'a context where there is a different criticizer for each of the people involved is highly implausible'. In the case of (114), on the other hand, 'it is readily imaginable [...] that Person A [...] has a person (Person X) who loves him/her, B has a person (Person Y) who loves him/her, [...] and so on' (Kuno & Takami 2002: 222). The alleged pragmatic implausibility of an inverse scope interpretation of (116) overrules the 'ambiguous' verdict of the expert system, rendering the sentence unambiguous. This is clearly not a convincing account, since scenarios *can* readily be found in which a distributive interpretation of *daremo* is highly plausible: a panel discussion, for example, where each member of the panel is criticised by another member of the panel or by a member of the audience.[75]

Quantitative data on QP-QP judgements are essential, here, to establish what the facts really are. The quantitative data on Japanese [∃-NOM...∀-ACC...] sentences collected from over 60 native Japanese informants during the course of the experimental work for the present dissertation (see Chapters 4 and 5) do not support Kuno *et al.*'s claim of the fundamental ambiguity of these sentences in Japanese. Most informants rejected inverse scope, as predicted by Hoji (1985), even though pragmatically plausible contexts for inverse scope interpretation were provided by means of pictures.

Kuno *et al.*'s account thus seems to be flawed due to predictions of scope ambiguity that are not supported by the experimental data. In addition, the framework seems to lack principled restraint, since it contains two mechanisms that can account for any judgement—individual or general—that does not correspond to what the expert system calculations predict. These are (i) the role of pragmatic (im)plausibility which can overrule any expert system verdict, and (ii) the potential for idiolectal variation in the relative weights of particular principles within an

[75] If there is a genuine difference between (114) and (116) in the availability of inverse scope, a more promising explanation might be found by considering the thematic roles of the arguments. The subject of the stative verb *aisuru* 'love' could be argued to have the θ-role experiencer, while the subject of the action verb *semeru* 'criticise' is a thematic agent. The object in each case is a theme. According to the thematic hierarchy, agents are more prominent than experiencers, which are more prominent than themes. Hence it follows that inverse scope of a theme over an experiencer (as in (114)) may be more readily available than inverse scope of a theme over an agent (as in (116)).

individual's expert system. In other words, the system is descriptive rather than explanatory: it can describe any variation of quantifier scope judgement.

Despite these drawbacks, Kuno *et al.*'s data are interesting in that they discuss a far wider variety of Japanese quantifiers than Hoji did.[76] The data also highlight the potential for variation in judgements about quantifier scope, and the framework itself attempts to account for this. In particular, Kuno *et al.* draw attention to the role of pragmatics and extra-syntactic mechanisms in scope interpretation, and they appeal to a view of scope determination arising from a number of factors. This view is compatible with parallel-processing accounts of syntactic ambiguity resolution, such as the competition model of MacWhinney (1987) and MacWhinney & Bates (1989), in which (as summarised by Kurtzman & MacDonald 1993: 274) any principles relevant to the sentence in question 'compete' in the building of its representation. If some principles favour one interpretation while others favour another, two representations may be built, and the final choice potentially selected randomly. On the other hand, if, during the building process, more principles favour one representation than another, only the most favoured representation is built. Kurtzman & MacDonald's (1993) empirical investigation of QP-QP scope interpretation detailed in the following chapter appeals to parallel-processing accounts of scope ambiguity resolution.

2.3.2.2. A pragmatic account of inverse scope (Hayashishita 1999, 2000a, 2000b)

Hayashishita's (1999, 2000a, 2000b) account of scope phenomena also makes reference to the lexical properties of different quantifiers. He identifies two types of quantifier (drawing on Liu 1990): those that can refer to a specific group (Type A, in (117a)) and those that cannot (Type B, in (117b)) (Hayashishita 2000b: 207).[77]

117. a. *Type A quantifiers:* QPs that can refer to a specific group
dareka 'someone'
subete-no (kaisya) 'all (the companies)'
sannnin-no (otoko) 'three (men)'

[76] In addition to *dareka* 'someone' and *daremo* 'everyone', Kuno & Takami (2002) investigate numeral quantifiers, and a variety of universal quantifiers.
[77] Hayashishita (2000b: 207 fn. 8) points out that Type B quantifiers can be made to refer to a specific group in an appropriate pragmatic context. For example, *three or more students* could refer to a specific group known to contain three or more students. Generally, however, this referring interpretation of Type B quantifiers is unavailable.

b. *Type B quantifiers:* QPs that cannot refer to a specific group
sannin-izyoo-no (gakusei) 'three or more (students)'
sukunakutomo sannin-no (gakusei) 'at least three (students)'
kanari-no-kazu-no (gakusei) 'a good number of (students)'

Only the Type A quantifiers can induce an object-wide distributive reading (i.e.. inverse scope) in [S...O...] word order. This reading is unavailable when the object is of Type B. Hayashishita claims that this generalisation holds for both English and Japanese. It is illustrated for Japanese in (118), where the object is a Type A QP, and (119), where the object is a Type B QP (examples based on Hayashishita 1999: 203):

118. (sukunakutomo) ippon-no ya-ga itutu-no mato-ni sasotta.
(at least) one-GEN arrow-NOM five-GEN target-DAT pierced
'At least one arrow pierced five targets.'
Interpretation: S>O; O>S

119. (sukunakutomo) ippon-no ya-ga itutu-izyoo-no mato-ni sasotta.
(at least) one-GEN arrow-NOM five-more.than-GEN target-DAT pierced
'At least one arrow pierced more than five targets.'
Interpretation: S>O; *O>S

Thus, for Hayashishita, there is, in principle, no difference between Japanese and English with regard to the availability of inverse scope in [S...O...] sentences. However, there *is* a difference between inverse scope and linear scope with regard to how the two types of scope interpretation are derived. Hayashishita argues that subject-wide distributive scope is derived directly from the LF representation, while object-wide distributive scope (in non-scrambled sentences) is derived from a post-LF cognitive representation.

The LF representation of a doubly-quantified sentence is assumed to retain the surface order of the two QPs. In other words, Hayashishita adopts a Scope Principle similar to that of Huang (1982) (see footnote 25) and Reinhart (1976):[78]

120. *Scope Principle* (Hayashishita 2000a: 289)
An NP α takes scope over an NP β only if α and β are in A-positions and α c-commands β at LF.

[78] Hayashishita's account leaves aside the question of whether Quantifier Raising occurs. If it does, quantifiers are raised to positions that reflect their surface order, and thus accord with the Scope Principle in (120).

The subject-wide scope interpretations of (118) and (119) are thus derived from the LF representations of the two sentences, as per the Scope Principle (120).

For the object-wide scope reading of (118), Hayashishita proposes a post-LF process of 'Subject Predication'. This is available in addition to the LF interpretation when the object QP is of Type A. Hayashishita's Subject Predication Hypothesis is presented in (121):

121. *Subject Predication Hypothesis* (Hayashishita 2000a: 290)
If a given sentence has the following LF representation, then at some cognitive level the 'value' of α can be interpreted as the Subject of a Predicate *S*.

S
... α ...
LF

$\Rightarrow$ [the 'value' of α] $\forall x\ (x \in X)$ $\lambda y\ [s \ldots y \ldots]\ (x)$

Subject Distributor Predicate

Hayashishita explains that Subject Predication proceeds in two steps (Hayashishita 2000a: 290):

> First, from a given LF representation, a Predicate is formed by substituting a variable for an NP α whose 'value' is to be the Subject. Second, some set γ is taken from the domain of the speaker's direct experience [...] to be the Subject of the Predicate, and γ must be able to be 'associated with' α.

Logically, the set γ can only be associated with the NP α (the object, in a sentence like (118)), if α is a QP that can refer to a specific set—in other words, a QP of Type A in (117a). Subject Predication, and hence object-wide distributive scope, is therefore not available for Type B QPs, since they cannot refer to a specific set.

Subject Predication requires that the person interpreting the sentence can actually imagine a relevant set with which the object QP can be associated. This may account for some of the individual differences in judgements about inverse scope: one individual may easily be able to imagine a relevant set, while the other cannot. However, a drawback of Hayashishita's analysis is that it leaves unanswered the question of why object-wide distributive scope should be so much more readily available with English *every* than with Japanese *wh*+QPt universals (both Type A

QPs).[79] If, as in Hayashishita's account, object-wide scope is always potentially available with Type A QPs via Subject Predication, something must prevent Subject Predication from occurring in Japanese. Similarly, Subject Predication must presumably be blocked in English when the object QP is *all the N*, despite this being among the Type A QPs.

Since Subject Predication is an extra-syntactic process, some light may be shed on what makes it less available in Japanese by considering other extra-syntactic accounts of scope interpretation. Working within Topic-Focus pragmatics, Erteschik-Shir (1997, 1999) argues that elements taking wide scope must be (pragmatic) topics within the discourse. In other words, they must refer to old information. Tomioka (2004) argues that Japanese *wh*+QPt quantifiers such as *daremo* 'everyone' or *dono N-mo* 'every N', along with disjunctive NPs, and even nominative-marked NPs, are 'Anti-Topic Items' (ATIs): items that cannot be construed as discourse topics. Each ATI has its own semantic or pragmatic reasons for being unable to act as a topic. For *wh*+QPt universals, Tomioka (2004: 8) suggests that their property of 'domain-widening', proposed by Kawashima (1994a), may rule out topic-hood. Domain-widening is a property also of English *any* (but, crucially, not of *every*), which widens the domain of a previously mentioned NP to include elements not originally assumed to be part of that NP. In other words, *wh+mo* QPs can admit new elements to the discourse, and this bars them from being the topic (i.e., old information). The following example shows how *daremo* 'everyone' can act as a domain-widener (Kawashima 1994: 129–130 (30)):

122. A: Gakusei-wa yob-ana-katta.
student-TOP invite-NEG-PAST
'(I) didn't invite any students.'

B: John (who is a student)-wa yon-da desyoo.
John-TOP invite-PAST right
'You invited John (who is a student), right?'

A: Iya, dare-mo yob-ana-katta.
no, who-MO invite-NEG-PAST
'No, (as for students) (I) didn't invite anyone.'

[79] The experimental results of this dissertation confirm that English *every* is judged as allowing object-wide scope whereas Japanese *dono...mo* 'every' is not. See Chapters 4 and 5.

The word *gakusei* 'students' in speaker A's first utterance in (122) specifies a domain of students. Speaker B's utterance shows that he understands this domain to exclude John even though John is a student. Speaker A's second utterance uses *daremo* 'everyone' to widen the domain of *gakusei* so that it includes everyone who is a student (i.e., including John, contra speaker B's narrower understanding of the domain of *gakusei*).

Before applying the notion of ATI status to *wh+mo* QPs in Hayashishita's framework, it is instructive to examine how Tomioka uses the concept in *wh*-QP questions. His goal is to account for Hoji's (1985) observation, which he endorses, that canonical [QP-subject ...*wh*-object ...] questions in Japanese ((18c), repeated in (123a)) are lower in acceptability than scrambled [*wh*-object ... QP-subject ...] questions ((18d), repeated in (123b)).

123. a ?? Daremo-ga nani-o kaimasita ka.
everyone-NOM what-ACC bought Q

b. Nani$_i$-o daremo-ga t_i kaimasita ka.
what-ACC everyone-NOM bought Q
'What did everyone buy? (*scrambled*)'

Tomioka notes that, in *wh*-questions, the non-*wh*-element is assumed to be old information in the discourse (Prince 1981; Schwarzschild 1999). In other words, it should contain the topic. Thus, if *wh+mo* QPs are ATIs, the unacceptability of (123a) is accounted for at the level of information structure: *daremo* 'everyone' is the non-*wh* part of the question, yet it cannot serve as the topic, and thus the question fails. Under this account, it is the acceptability of (123b) that is problematic. Tomioka (2004: 9–10) solves the problem by appealing to the phonological effect of scrambling on information structure. He argues that scrambling of the *wh*-element means that the post-*wh* part of the sentence becomes de-accented, or down-stepped. The down-stepped part of a sentence is associated with old discourse information. In the framework of Vallduví (1992, 1995), old information (the *ground*) consists of a *link* and a *tail*. Tomioka assumes that the link is the topic-marked part of the sentence and, therefore, cannot be a *wh*+QPt QP since these are ATIs. However, in information structure, the down-stepped *wh*+QPt can become part of the tail and hence, part of the old information in the question. Thus, (123b) is legitimised despite the fact that *daremo* 'everyone' cannot serve as the topic. Phonological downstepping of *daremo* is not possible in (123a) because of *daremo*'s sentence-initial position, which is an unnatural site for the flat low pitch of a down-stepped item. The difference in acceptability between (123a) and (123b) is thus a question of pragmatics and phonology. Tomioka (2004: 11) points out that this makes it unsurprising that judgements about the acceptability of canonical [*wh*-object ... QP-subject ...] questions like (123a) should be fragile: '[h]ow a speaker judges intervention effects [i.e., the prohibition against a *wh*+QPt QP c-commanding a *wh*-phrase] depends on how accommodating the speaker can be in dealing with pragmatic difficulties caused by less-than-perfect information structure'.

Returning to Hayashishita's account of QP-QP interactions, the relevance of Tomioka's (2004) proposal is that it shows how information-structure factors are implicated in the behaviour of QPs. In addition, Tomioka reaches a conclusion similar to that which can be drawn from Hayashishita's framework about what may account for native speaker variability with respect to QPs, namely that individual responses to pragmatic difficulties in QP interpretation may play a role. Combining Erteschik-Shir's (1997, 1999) proposal that only information-structure topics can take wide scope, and Tomioka's proposal that *wh*+QPt QPs are ATIs, the lack of application of Subject Predication (and hence, object-wide scope) to these QPs is predicted, since the Subject in Subject Predication must essentially be a topic, being defined (as previously cited) as 'some set [...] from the domain of the speaker's direct experience' (Hayashishita 2000a: 290). Since subject-wide scope is available with *wh*+QPt QPs, however, their ATI status must be able to be softened in certain cases. This could plausibly be achieved by manipulation of information-structure factors: Tomioka (2004: 23, fn. 4) notes, for example, that when *wh*+QPt existentials, such as *dareka* 'someone', receive contrastive stress, they can be topics. Alternatively, Hayashishita's Scope Principle (120) offers a syntactic means for subject QPs (in canonical sentences) to take wide scope, by virtue of c-commanding the object.

The above discussion is intended only to suggest a means by which information-structure accounts of QP interpretation might expand Hayashishita's framework so that it can account for the lack of availability of object-wide scope for his Type A quantifiers in Japanese. Many problems remain. The lack of object-wide scope for *subete-no N*/*all the N* in Japanese and English, for example, is unaddressed.[80] Moreover, the accounts of both Erteschik-Shir (1997, 1999) and Tomioka (2004) do not make use of LF structure, while Hayashishita refers specifically to LF. The accounts thus cannot simply be amalgamated. However, exploration of information-structure processes is clearly a promising direction to take in attempting to account for the restrictions on Hayashishita's essentially pragmatic account of object-wide scope.

2.3.2.3. Comment

As noted in Section 2.3.2.1, the fact that the functional account of scope resolution by Kuno (1991) Kuno & Takami (2002), and Kuno *et al.* (1999, 2000) appears to make wrong predictions about scope judgements in Japanese means that the implications of that account for L2 acquisition are not explored. This section focuses on the implications of the account by Hayashishita (1999, 2000a, 2000b).

Under Hayashishita (1999, 2000a, 2000b) subject-wide scope is a syntactic phenomenon, while object-wide scope arises via the extra-syntactic process of Subject Predication. Accordingly, the cross-linguistic differences in the availability of object-wide scope (i.e., O>S scope is readily available in English, but not in Japanese, Korean or Chinese [in non-scrambled contexts]) may be due to the inability of Japanese, Korean and Chinese QPs to undergo Subject Predication. For Japanese, it was proposed in the previous section that the fact that *wh*+QPt QPs cannot readily act as topics may play a role in their apparent inability to undergo Subject Predication. If this is correct, part of the acquisition problem for English learners of Japanese entails coming to know that these Japanese *wh*+QPt quantifiers cannot be topics, while in English, *every N* and *some N* can. English learners of Japanese are thus predicted to, at least initially, allow non-target-like inverse scope in Japanese due to transfer of the lexical properties of English quantifiers to their interlanguage. Restructuring of the interlanguage could come about by exposure to the relevant Japanese quantifiers in contexts that make the ATI status of these elements clear. An example of such a context would be that given in (122), which showed the 'domain-widening' property of *daremo* 'everyone'. The property of domain-widening is argued to be incompatible with the property of topic-hood. It seems that examples like (122), which unambiguously demonstrate the ATI

[80] Indeed, Tomioka (2004) specifies that *subete-no N* 'all the N' is not an ATI: it can act as a topic, and, unlike *wh*+QPt universals, it is perfectly acceptable in a position that c-commands a *wh*-phrase.

properties of *wh*+QPt quantifiers, are unlikely to occur frequently in the input. They are also not a topic of L2 instruction. Thus, the L1 English-L2 Japanese interlanguage is unlikely to undergo the relevant restructuring at an early stage of L2 acquisition.[81]

Korean also has *wh*+QPt quantifiers. Tomioka (2004) argues that, like their Japanese counterparts, these are ATIs. Thus, for Korean learners of Japanese, there is no acquisition problem. However, for Chinese, It is less clear what might contribute towards ruling out Subject Predication (and, consequently, object-wide scope). One possibility relates to the observation that post-verbal NPs in Chinese tend to be indefinite (Liu 1997, among others),[82] and consequently not topics. However, this may amount to the same thing as holds true for English (and other languages): that topics tend to occur closer to the beginning of a sentence, probably because, in the words of Erteschik-Shir (1999: 130), 'processing is facilitated when the topic ("what we are talking about") precedes the predicate ("what we say about the topic")'. Thus it does not necessarily account for why object-wide scope should be less available in Chinese than in English. However, since ATI status in Japanese and Korean is argued to arise from the specific semantic or pragmatic properties of individual ATIs, Chinese quantifiers may conceivably also have specific properties that differentiate them from English *every/some* with regard to object-wide scope. If these properties transfer, then Chinese learners of Japanese are predicted to exhibit target-like rejection of inverse scope in Japanese from the outset.

To summarise, the difference in the availability of object-wide scope between English, on the one hand, and Japanese, Korean and Chinese, on the other, may arise, under Hayashishita (1999, 2000a, 2000b), because certain quantifiers in Japanese, Korean and Chinese have properties rendering them less likely to act as discourse topics than English quantifiers. If this is correct, then the acquisition problem for English learners of Japanese includes coming to know the relevant lexical properties of Japanese quantifiers that exclude them from acting as topics. By contrast,

[81] Note that the fact that *wh*+QPt quantifiers are incompatible with the Japanese topic-marker *wa* will not provide direct evidence to learners of the ATI status of these quantifiers. This is because failure to encounter *wh*+QPt quantifiers marked with *wa* does not preclude that such quantifiers should ever be marked with *wa*, given that the L1-(English)-based interlanguage allows all quantifiers to behave as topics.

[82] Liu (1997) emphasises that this is only a tendency, and notes (Liu 1997: 87) that 'in terms of well-formedness, virtually any NP is allowed post-verbally'. However, some sentences may be improved by moving a definite NP to a pre-verbal position.

quantifiers in Chinese-Japanese interlanguage and Korean-Japanese interlanguage should lack the topic-hood properties from the outset, due to L1 transfer. Thus, for Chinese-speaking and Korean-speaking learners of Japanese, the lack of object-wide scope in Japanese SOV QP-QP sentences should not present an acquisition problem.

2.4. Conclusion

This chapter began by outlining some key features of quantifier scope interpretation in Japanese, English, Chinese and Korean. It was noted that (i) English allows object-wide scope in doubly-quantified [S...O...] sentences, while Japanese, Chinese and Korean do not;[83] and (ii) English and Chinese allow pair-list readings in questions with a QP-subject and *wh*-object, while Japanese and Korean do not.[84] Two L2 poverty-of-the-stimulus problems were identified in the acquisition of the Japanese scope phenomena by native English-speaking and native Chinese-speaking learners, as summarised in (124).

124. *L2 poverty of the stimulus phenomena investigated in this dissertation*:
   a. acquisition of the lack of inverse scope in Japanese SOV QP-QP sentences by English-speaking learners of Japanese
   b. acquisition of the lack of pair-list readings in scrambled Japanese [*wh*-object ... QP-subject ...] questions by both English-speaking and Chinese-speaking learners of Japanese.

These are L2 poverty-of-the-stimulus problems because (as detailed in Section 2.1) the relevant facts of Japanese are underdetermined by the input, by the learners' L1 knowledge, and by classroom instruction, which does not cover quantifier scope interpretation.

Sections 2.2 and 2.3 presented a number of theoretical accounts of quantifier scope determination, focusing in particular on the cross-linguistic differences between English, Japanese, Chinese, and Korean. Five analyses that account for scope determination in terms of syntax were detailed first: May (1977, 1985); Hoji (1985); Aoun & Li (1993); Hornstein (1995); and Beghelli (1995, 1997); Beghelli & Stowell (1997). Section 2.3.2 then outlined analyses that additionally appeal to extra-syntactic factors, including the semantic properties of quantifiers, and the role of

[83] Although, as described, scrambled QP-QP sentences are ambiguous in Japanese and Korean.
[84] As already noted (footnote 19), the claim that Korean lacks pair-list interpretations in scrambled [*wh*-object...QP-subject...] questions is challenged in Chapter 5 on the basis of quantitative experimental native Korean data.

pragmatics. Each analysis was shown to have advantages and disadvantages, with no single framework accounting unproblematically for all the scope phenomena discussed.

Within the L2 acquisition framework of this dissertation (described in Chapter 1), it is not a problem that the precise determinants of scope interpretation remain somewhat elusive. This is because of the dissertation's focus on L2 acquisition under poverty of the stimulus. Whatever precise mechanisms account for the Japanese scope phenomena, the L2 poverty-of-the-stimulus problems identified in (124) remain. In other words, if learners demonstrate native-like knowledge of the Japanese scope phenomena, this can be taken as evidence that their L2 acquisition of quantifier scope is constrained by the same mechanisms constraining native Japanese, whatever those mechanisms are.

Predictions about how the poverty-of-the-stimulus problems listed in (124) might be overcome—or whether they can be overcome at all—differ depending on the particular account of the relevant phenomena, as outlined for each account throughout the chapter. For (124a)—acquisition of the lack of inverse scope in Japanese SOV QP-QP sentences by English-speaking learners of Japanese—the possible 'solution' to the problem under each account is summarised in (125), along with details of the evidence required to 'trigger' the solution, and predictions about how acquisition might proceed.

125. *Acquisition of the lack of inverse scope in Japanese SOV QP-QP sentences by English-speaking learners of Japanese*

a. *Account*: Hoji (1985)
   Solution: Learners must reset a scope interpretation parameter from its 'ambiguous' setting in English to its 'unambiguous' setting in Japanese.
   Evidence: Logically, there can be no evidence in the input to indicate the unambiguity of Japanese, given that the L1-based interlanguage allows for ambiguity.
   Prediction: Learners will be unable to reset the parameter and hence unable to acquire native-like QP-QP scope interpretation.

b. *Account*: Aoun & Li (1993)
   Solution: Learners must reset a VP-internal subject parameter from its 'VP-to-IP subject-raising' setting to its 'no VP-to-IP subject-raising' setting'. Knowledge of the lack of O>S scope in Japanese SOV QP-QP sentences will arise automatically from this resetting.

	Evidence:	Empirical evidence to motivate resetting of a VP-internal subject parameter is likely to be unclear given that the evidence to motivate adoption of a 'no VP-to-IP subject-raising' account of Japanese is inconclusive. Even if Japanese really does lack VP-to-IP subject-raising, the L1-based interlanguage grammar, in which subjects raise to IP, can (arguably) represent Japanese adequately—if incorrectly.
	Prediction:	Learners will be unable to reset the parameter and hence unable to acquire native-like QP-QP scope interpretation.
c.	*Account*:	Hornstein (1995)
	Solution:	Learners must reset a VP-internal subject parameter, as for Aoun & Li (1993).
	Evidence:	As for Aoun & Li (1993), in (125b), above.
	Prediction:	As for Aoun & Li (1993), in (125b), above.
d.	*Account*:	Beghelli (1995, 1997); Beghelli & Stowell (1997)
	Solution:	Learners must (possibly) come to know that there is no syntactic singular/plural distinction in Japanese. Knowledge of this lack of singular/plural distinction will automatically rule out O>S scope in Japanese SOV QP-QP sentences.
	Evidence:	Japanese does not have compulsory plural marking. From the point of view of the English-based interlanguage, a plural context is a 'required context' for plural morphology on NPs. Thus, the lack of plural marking in such required contexts in Japanese may constitute the necessary indirect negative evidence to trigger restructuring of the interlanguage grammar to one that lacks a singular/plural distinction. However, the clarity of this negative evidence will be blurred by the facts that (i) optional plural marking can be used on human nouns in Japanese, and (ii) the English-based interlanguage can parse nouns without any singular or plural morphology as if they are mass nouns.
	Prediction:	Acquisition of native-like scope interpretation is possible. However, impediments to the transparency of the evidence may mean that lower-proficiency learners allow non-native-like object-wide scope.
e.	*Account*:	Hayashishita (1999, 2000a, 2000b)[85]
	Solution:	Learners must (possibly) come to know that Japanese *wh*+QPt quantifiers cannot be construed as discourse topics. Knowledge of the incompatibility of *wh*+QPt quantifiers with topic-hood will automatically rule out O>S scope in Japanese SOV QP-QP sentences.

[85] Recall that Hayashishita's (1999, 2000a, 2000b) account does not actually differentiate between English and Japanese with respect to QP-QP scope interpretation. The proposal referred to here, whereby the ability of a universal quantifier to be a topic affects its ability to take wide scope, draws on the hypotheses of Erteshik-Shir (1997, 1999) and Tomioka (2004), within the framework of Hayashishita's account.

Evidence: Whether or not a QP can be a discourse topic is argued to be a consequence of its semantic or pragmatic properties. Encounters with quantifiers in contexts that unambiguously indicate the relevant semantic or pragmatic properties could thus motivate IL restructuring. For example, encountering a *wh*+QPt quantifier in a context that indicates its domain-widening property (see (122)) may lead to that quantifier being ruled out from acting as a topic, domain-widening being incompatible with topic-hood.

Prediction: Acquisition of native-like scope interpretation is possible. However, it is likely to proceed slowly since occurrences of quantifiers in contexts that unambiguously indicate their lack of topic-hood are rare.

In short, English-learners of Japanese are predicted to be unable to acquire scope rigidity in Japanese SOV QP-QP sentences under the accounts in (125a–c) (Hoji 1985; Aoun & Li 1993; and Hornstein 1995). However, under the accounts in (125d–e) (the Target Landing Sites theory by Beghelli (1995, 1997) and Beghelli & Stowell (1997); and Hayashishita's (1999, 2000a, 2000b) account), acquisition is predicted to be possible but slow. In other words, under these two accounts (125d–e), higher proficiency English-speaking learners of Japanese are predicted to demonstrate knowledge of the lack of object-wide scope in Japanese SOV QP-QP sentences; however, lower proficiency English-speaking learners of Japanese are not. Under all five accounts in (125), English-speaking learners of Japanese will differ from Chinese-speaking and Korean-speaking learners of Japanese, since the latter two learner populations will demonstrate knowledge of the lack of object-wide scope in Japanese SOV QP-QP sentences from the outset (i.e., even at lower proficiency levels), due to L1 transfer.

Turning to the second poverty-of-the-stimulus problem (124b)—acquisition of the lack of pair-list readings in scrambled Japanese [*wh*-object ... QP-subject ...] questions by English-speaking and Chinese-speaking learners of Japanese—only the analysis by Aoun & Li (1993) provided a satisfactory account of the Japanese data and the Chinese and English data. Item (126) presents the possible solution to the poverty-of-the-stimulus problem under Aoun & Li (1993), along with details of the evidence required to 'trigger' the solution, and a prediction about how acquisition might proceed.

126. *Acquisition of the lack of pair-list readings in scrambled Japanese [wh-object...QP-subject...] questions by English-speaking and Chinese-speaking learners of Japanese*

Account: Aoun & Li (1993)

Solution: Learners must come to know that (covert) *wh*-movement in Japanese is QR, not *wh*-movement.

Evidence: *Wh*-movement as QR may be linked to the existence of question marking particles (e.g., *ka*) in Japanese. The frequent and clear occurrence of question markers in the Japanese input could plausibly motivate the relevant interlanguage restructuring.

Prediction: Learners will be able to acquire native-like [*wh*-object...QP-subject...] question interpretation.

Thus, under Aoun & Li's account, Chinese-speaking and English-speaking learners of Japanese are predicted to be able to acquire the lack of pair-list readings in scrambled Japanese [*wh*-object...QP-subject...] questions, despite poverty of the stimulus. Nonetheless, these two learner populations are expected to differ from Korean-speaking learners of Japanese at lower levels of proficiency, since Korean-speaking learners of Japanese should have knowledge of the lack of pair-list readings from the outset, due to L1 transfer.[86] For Chinese-speaking and English-speaking learners, acquisition of the relevant knowledge will take some time. Thus, at lower levels of proficiency, these learners may accept non-native-like pair-list interpretations of Japanese [*wh*-object...QP-subject...] questions, due to the interlanguage grammar not yet having undergone the necessary restructuring.

The experimental work conducted for this dissertation aims to investigate the predictions, detailed above, about L2 knowledge of Japanese quantifier scope interpretation by English-speaking, Chinese-speaking and Korean-speaking learners. The main goal of the research is, as outlined in Chapter 1, to increase our knowledge of the roles of UG and L1 transfer in L2 acquisition. However, the experimental results may also be informative with respect to which theoretical account (or accounts) of quantifier scope interpretation is likely to be correct. For Japanese QP-QP interpretation, if English-speaking learners can overcome poverty of the stimulus and acquire the lack of object-wide scope in Japanese SOV QP-QP sentences, this could be evidence in support of the accounts by Beghelli (1995, 1997), Beghelli &

[86] Predictions about Korean-speaking learners of Japanese with respect to pair-list interpretations of Japanese [*wh*-object...QP-subject...] questions are revised in Chapter 5 on the basis of native Korean data (mentioned in footnotes 19 and 86) indicating that pair-list interpretations are possible in scrambled Korean [*wh*-object...QP-subject...] questions, contra the claims of the theoretical literature.

Stowell (1997) and Hayashishita (1999, 2000a, 2000b), since the other accounts (125a–c) predict that the lack of object-wide scope cannot be acquired. For the interpretation of Japanese [*wh*-object...QP-subject...] questions, if advanced English-speaking and Chinese-speaking learners demonstrate knowledge of the unavailability of a pair-list interpretation, this could be evidence in support of Aoun & Li's (1993) account.

A further utility of the study is to provide experimental data on native knowledge of QP-QP and *wh*-QP interpretation in English, Japanese, Chinese and Korean. This is particularly crucial for Japanese and Korean, as there is currently very little experimental data available for these languages.[87] An important observation that has emerged during the course of this chapter is that judgements about scope interpretation are subject to a large degree of individual variability. The present investigation of native judgements about scope interpretation—as control data for comparison with the L2 judgements—will thus contribute towards identifying which generalisations are robust under experimental conditions.

[87] To the author's knowledge, there are no quantitative studies of Korean quantifier scope interpretation, and only one of Japanese QP-QP interpretation: Sano (2004) investigates the judgements of ten adult native Japanese speakers about one QP-QP sentence. See Chapter 3.

3. Investigating quantifier scope interpretation: Previous studies

3.0. Introduction

The body of psycholinguistic research on scope interpretation of quantificational determiners in L2 acquisition is still relatively small. This chapter reviews four studies, which, to the best of the author's knowledge, constitute the key existing studies of this topic: Lee, Yip & Wang (1999a) investigate QP-QP interactions in L1 Chinese-L2 English interlanguage; while Miyamoto & Yamane (1996), Miyamoto & Takata (1998) and the study reported in Dekydtspotter, Sprouse, Swanson & Thyre (1999) and Dekydtspotter, Sprouse & Swanson (2001) investigate *wh*-QP interactions, the former two in L1 Japanese-L2 English interlanguage, the latter in L1 English-L2 French interlanguage.[88] In addition, two studies of adult native-speaker knowledge of QP-QP scope interpretation are presented: Kurtzman & MacDonald (1993), focusing on English, and Lee, Yip & Wang (1999b), investigating English and Chinese. Finally, a study by Sano (2004) is presented, which investigates knowledge of QP-QP interpretation in the L1 acquisition of Japanese, and includes data on adult native Japanese judgements about one QP-QP sentence. To the author's knowledge, there are no existing experimental studies of knowledge of quantifier scope interpretation in native Korean.[89]

The studies of adult native-speaker knowledge of QP-QP scope are presented first, in Section 3.1. Section 3.2 presents the L2 acquisition studies. For each study, the format adopted in Sections 3.1 and 3.2 is to first present details of the procedure, results, and conclusions, and then to comment on the study. Section 3.3 summarises the chapter.

[88] Some studies of L2 scope interpretation with quantifiers other than quantificational determiners exist: Dekydtspotter, Sprouse & Thyre (1998) and Clark-Peasey (1999) investigate quantificational adverbs; Haegeman (1985) looks at the scope of negation and modal verbs. In addition, Outcalt & Dekydtspotter (2004) expand on the study reported by Dekydtspotter *et al.* (1999) and Dekydtspotter *et al.* (2001), to investigate the role of context in non-native scope disambiguation.

[89] For an overview of investigations into quantifier scope interpretation in L1 acquisition, see Guasti (2002, Chapter 9). As Guasti notes (2002: 313), work in this field, too, is just beginning. Experimental studies indicate that knowledge of universal quantifiers is acquired by age 4–5, although results are inconclusive as to whether, by that stage, knowledge is fully adult-like. Since the present study investigates only adult knowledge, further details of quantifier scope interpretation in L1 acquisition are omitted here.

3.1. Native knowledge of quantifier scope interpretation

3.1.1. Kurtzman & MacDonald (1993)

The aim of Kurtzman & MacDonald's (1993) study (hereafter K&M, in this sub-section) is to investigate the role of structural position in the disambiguation of doubly-quantified sentences in English. They investigate the interaction between *every* and *a* with action verbs and verbs of perception, and in active sentences and passive sentences.[90] The sentence types investigated are as follows:

127. *Active sentence types*
   a. *'Every ...a' order, action verb*: Every kid climbed a tree.
   b. *'A ...every' order, action verb*: A kid climbed every tree.
   c. *'Every ...a' order, perception verb*: Every agent observed a spy.
   d. *'A ...every' order, perception verb*: A spy observed every agent.

128. *Passive sentence types*
   a. *'Every ...a' order, action verb*: Every tree was climbed by a kid.
   b. *'A ...every' order, action verb*: A tree was climbed by every kid.
   c. *'Every ...a' order, perception verb*: Every spy was observed by an agent.
   d. *'A ...every' order, perception verb*: A spy was observed by every agent.

All the sentences in (127) and (128) are potentially ambiguous: either the first or the second QP can take wide scope. K&M aim to find out whether or not the preferred interpretation favours wide scope of the first QP. The active and passive sentence types were investigated in two separate experiments with the same experiment design. For each sentence type, four two-sentence pairs were constructed, each with 16 versions using different nouns and verbs.[91] Examples for (127a) are given in (129a–d) (K&M: 252):

129.

	Test sentence	Continuation with wide scope of QP1	Continuation with wide scope of QP2
a.	Every kid climbed a tree.	The trees were full of apples.	(n/a)
b.	Every kid climbed a tree.	(n/a)	The tree was full of apples.
c.	Every kid climbed a different tree.	The trees were full of apples.	(n/a)
d.	Every kid climbed the same tree.	(n/a)	The tree was full of apples.

[90] In separate experiments, K&M also investigate scope interpretation in complex NP constructions, as in *George has [a photograph of every admiral]*. Details of these experiments are not reported here.

[91] Note that type (127a) uses the same nouns and verbs as type (127b), and type (127c) the same as type (127d), so the total number of different noun and verb combinations is 32 (16 x 2) (rather than 64 (16 x 4)). Similarly, the passive sentence set (with types as given in (128)) comprises 32 versions. See K&M appendices for full details of the test sets.

The sentence in (129a–b) is ambiguous: *every kid* could take wide scope ('for every kid there is a tree that she/he climbed'), in which case there must be more than one tree, as in the continuation in (129a); or *a tree* could have wide scope ('there is a tree, such that every kid climbed that tree'), in which case there is only one tree, as in the continuation in (129b). The sentences in (129c–d), on the other hand, are unambiguous: the object *a different tree* in (129c) forces a plural construal of *tree*, so only the continuation (129c) with the plural *the trees* is possible (wide scope of QP1); *the same tree* as object in (129d) forces a singular construal of *tree*, so only the continuation (129d) with the singular *the tree* is possible (wide scope of QP2). The unambiguous sentences were used as control items to compare with the ambiguous sentences.

The test stimuli (256 for each experiment) were divided into eight sets of 32. Each set was combined with 88 filler pairs, 40 of which had continuation sentences that were compatible with the first sentence, and 48 of which had incompatible continuation sentences. Twenty practice items were added to the beginning of each test set. Thus eight different test files were created for each experiment.

The test participants were two groups of 48 native-English-speaking university students. One group took the active sentence test, the other, the passive sentence test. Each of the eight test files was used by six participants. Participants were tested individually using a computer terminal. The participant pressed the space bar to bring a test sentence onto the screen. After reading the sentence, touching the space bar brought the continuation onto the screen, and the participant used *Yes* and *No* buttons to indicate whether the continuation 'made sense' and was a 'natural continuation' of the first sentence (K&M: 255). Participants were encouraged to read the sentences quickly but to ensure that they understood them. Each test session lasted approximately 30 minutes.

K&M found that, with both actives and passives, the continuations for the unambiguous test items (e.g., (129c–d)) were judged compatible with the test sentences at least 80% of the time,[92] regardless of the verb type (action or perception), the quantifier order ('*every*...*a*' or '*a* ...*every*'), or the location of wide scope (on the first or second QP). This rate is comparable with the 85% rate of

[92] Approximate percentages are reported here because they are taken from bar charts (K&M: 255, 260) that do not reveal the exact percentages.

correct selection of 'no' on the filler items with incompatible continuations, indicating that the participants were not answering randomly during the test. On the ambiguous sentences, specific patterns emerged, depending on the sentence type. On the ambiguous active sentences, acceptability (i.e., 'yes' judgements) of subject-wide scope (wide scope of the first QP) was above 70%, while acceptability of inverse scope (wide scope of the second QP) was between 20% and approximately 35% for all conditions except the *Every ...a* order with perception verbs (e.g., (127c)), where the rate was around 55%. Overall, wide scope of the first QP was significantly preferred over wide scope of the second QP (*min F'*(1,65) = 41.01, *p*<.001) on active sentences (K&M: 256). In addition, there was a significant effect for quantifier order on the inverse scope test items: there were more 'yes' judgements following the *Every...a* order than the *A ... every* order (*min F'*(1,75) = 12.07, p<.01) (K&M: 256). Quantifier order also interacted significantly with verb type on the inverse scope test items, due to the relatively high rate of 'yes' responses for inverse scope with the *Every ... a* order and perception verbs. On the ambiguous passive sentences, no robust effects were evident, yet acceptability rates for all conditions were lower than for the unambiguous sentences, ranging between 50% and 70%.

To summarise, the results showed a clear preference for subject-wide scope interpretation in the active QP-QP sentences, while in the passive QP-QP sentences, there was no clear preference for subject-wide scope or inverse scope. In addition, the inverse scope interpretation was more readily available in the active sentences with *a*, rather than *every*, as object quantifier, particularly with perception verbs.

K&M suggest that their results indicate the involvement of a number of principles affecting scope determination in scope disambiguation. They suggest that any or all of the following (hypothesised) principles could account for the overall preference for wide scope in active sentences (K&M: 246–8):[93]

130. a. a *linear order principle*: wide scope is assigned to the left-most QP (e.g., Kroch 1975; Lakoff 1971).
b. a *surface subject principle*: wide scope is assigned to the surface subject (Ioup 1975).
c. an *external subject principle*: wide scope is assigned to the external argument of the verb (based on Ioup 1975).

[93] All these hypothesised principles were introduced—if only briefly in some cases—in Chapter 2.

d. a *c-command principle*: wide scope is assigned to whichever QP c-commands the other at surface structure (Reinhart 1983. Also Hayashishita 1999, 2000a, 2000b; Huang 1982).[94]
e. a *topic principle*: wide scope is assigned to the semantic topic (Kempson & Cormack 1981; May 1985. Also Erteschik-Shir 1999).
f. a *thematic hierarchy principle*: wide scope is assigned to the agent in preference to the experiencer, and the experiencer in preference to the theme (Grimshaw 1990; Jackendoff 1972).

Although (130a–f) would all predict subject-wide scope in active sentences, principles like (130c) and (130f) predict wide-scope of the agent rather than the grammatical subject in passive sentences. K&M suggest that the lack of strong preference for (grammatical-) subject-wide scope or agent-wide scope in passives may be indicative of a parallel processing mechanism, whereby multiple interpretations of ambiguous sentences are built up in parallel. When some of the principles favour one interpretation but others favour another (as would happen with the principles in (130) applied to passives), the choice of interpretation may ultimately be random.

3.1.1.1. Comment

In the context of the present dissertation, the finding of particular interest is the relatively low rates of acceptance of object-wide scope on active QP-QP sentences, even though it is theoretically possible in English. This result could be due to the experiment design, whereby participants read the QP-QP sentence first, and then the disambiguating continuation was revealed. In such a design, the participants would already have formed an interpretation of the QP-QP sentence before reading the continuation. If the interpretation they have formed is the subject-wide-scope interpretation (as is likely, since this interpretation seems generally to be the most readily available), then the object-wide-scope continuation would clash with the interpretation already in mind, hence leading to participants to reject the object-wide-scope interpretation. In other words, the experimental procedure favours discovery of what interpretation of an ambiguous sentence is determined first, rather than discovery of which interpretations are actually possible. The focus of the present

[94] As detailed in Chapter 2, the analyses of May (1977, 1985), Hoji (1985), Aoun & Li (1993), Hornstein (1995), Beghelli (1995, 1997) and Beghelli & Stowell (1997), presented in Chapter 2, also predict wide scope for whichever QP c-commands the other, but in these analyses the c-command relation is at LF, and does not necessarily reflect surface order, due to the mechanism of QR.

research is to find out whether inverse scope is possible, in native English, Japanese, Korean and Chinese, as well as in non-native Japanese. K&M's results show that care must be taken to design an experiment that facilitates identification of the possibility of an interpretation, even if that interpretation may not be the first that comes to mind.

3.1.2. Lee, Yip & Wang (1999b)

Lee, Yip & Wang (1999b) (hereafter LY&Wb, in this section) investigate knowledge of scope interpretation in QP-QP sentences of native Chinese and English. They note that (as outlined in Chapter 2) Chinese is characterised as not allowing inverse scope in doubly-quantified sentences. This is shown in (131), repeated from Chapter 2.

131. Meige xuesheng dou mai-le yiben shu — unambiguous: S>O; *O>S
every student all buy-ASP one book
'Every student bought a book' (Huang 1982: 112, (3))

However, LY&Wb note the evidence from Beghelli & Stowell (1997), Ioup (1975), Jackendoff (1983), and others, that manipulation of the type of quantifier and the thematic role of the quantified object can influence the availability of inverse scope. Their investigation is designed to quantify these influences in Chinese and English.

The test instrument used was an evaluation task, completed by 30 educated adult native speakers of Mandarin Chinese and 27 educated adult native speakers of English. The participants were presented with doubly-quantified sentences of the types in (132) ('NumP' = a numerically quantified noun):

132.

	subject QP	object QP	verb type	example (LY&Wb: 180)
a.	NumP	*every N*	action	In this classroom, two boys cleaned every desk
b.	NumP	*all the N*	action	At this concert, two singers sang all the songs.
c.	NumP	NumP	action	Three workers repaired two machines.
d.	*every N*	NumP	action	In this classroom, every boy cleaned two desks.
e.	*all the N*	NumP	action	At this concert, all the singers sang two songs.
f.	NumP	*every N*	unaccusative	Two balloons floated to every roof in the area.
g.	NumP	*all the N*	unaccusative	Two flags waved from all the windows in the building.
h.	NumP	NumP	unaccusative	Two blankets hung on three clothes lines in the yard.

Manipulating the verb type enabled investigation of thematic role, with the subject and object taking agent and theme roles respectively in (132a–e), and theme and goal/location roles respectively in (132f–h). In the Chinese version of the test, *mei CL N* corresponded to *every N* and *suoyoude N dou* corresponded to *all the N*.

Four tokens of each type in (132) were constructed. Each token was presented to the participants accompanied by three interpretations, or four when both subject and object were numeral phrases, yielding a total of 100 interpretations to be judged. An example for type (132f) is provided in (133), with the types of interpretation indicated in brackets (LY&Wb: 181 (34)):

133. Two flags waved from every window in the building.
   A) There were only two flags. One flag waved from all the windows. The other flag also waved from all the windows. *(S>O)*[95]
   B) There were only two flags. One flag waved from some windows. The other flag also waved from the other windows. *(cumulative interpretation)*
   C) Each of the windows had two flags waving. Different windows had different flags. There were more than two flags. *(O>S)*

The test sentences and interpretations were presented in written questionnaire format. Participants rated each interpretation on a scale of 0 to 3, where '0' means '[the interpretation is] impossible and never understood that way'; '1' means 'possible but rarely understood that way'; '2' means 'possible and sometimes understood that way'; and '3' means 'possible and often understood that way'. This scale was chosen rather than a binary choice of 'possible'/'impossible' with the aim of preventing preferences about scope judgements from totally obscuring judgements of possibility: '[i]f asked to give a categorical yes-no judgment, subjects may reject marked readings which are in fact possible but highly dispreferred' (LY&Wb: 181).

In analysing the results, responses of '0' were considered to indicate rejection of the relevant interpretation, while '1', '2' and '3' were considered to indicate acceptance. Types 132c, d and e were used to investigate whether different quantifier types in subject position affected the availability of subject-wide scope. It was found that, in both Chinese and English, S>O scope was accepted over 85% of the time when the subject quantifier was a universal (*every/meige*, *all/suoyoude*), but with a

[95] LY&Wb refer to this interpretation as the 'each-all' interpretation.

NumP subject, the acceptance rate fell below 37% (LY&Wb: 182). The results showing acceptability of object-wide scope with different object quantifier types and roles (Types 132a, b, c, f, g, and h) are shown in Table 3 (based on LY&Wb: 183, 184).

Table 3: Acceptability of inverse scope in native English and native Chinese with different object quantifiers and object roles ('rating' = mean rating on scale of 0–3; '%' = rate of consistent acceptance)[a]

object QP	native English (n=27)				native Chinese (n=30)			
	object role: theme		goal/loc		*object role:* theme		goal/loc	
	rating	%	rating	%	rating	%	rating	%
every/meige	1.17	44.4	2.16	77.8	0.58	26.7	1.21	36.7
all/suoyoude	0.28	11.1	1.69	66.7	0.28	10.0	0.77	36.7
NumP	0.25	11.1	1.02	29.6	0.36	13.3	0.51	16.7

[a] 'Consistent acceptance' = acceptance of inverse scope on at least three of the four tokens for the relevant sentence type.

A number of patterns are evident in Table 3. First, a general difference is observable between English and Chinese: inverse scope is more readily available in English than in Chinese, with the highest rate of consistent acceptance being 77.8% in English (with *every* on the goal/location object condition), but only 36.7% in Chinese (with *meige* and *suoyoude* on the goal/location object condition). However, in both English and Chinese the type of quantifier and the object role seem to affect the availability of inverse scope. LY&Wb report significantly higher ratings for inverse scope with *every/meige* compared with *all/suoyoude* in the theme object condition in both English and Chinese. In English, inverse scope was also significantly more acceptable with *every* than with numeral quantifiers. In the goal/location condition, significantly higher ratings are reported for inverse scope in both languages with *every/meige* compared with *all/suoyoude* and numerals, and with *all/suoyoude* compared with numerals (LY&Wb: 183–4). To summarise, LY&Wb's results support the position that inverse scope is less readily available in Chinese than in English. However, they also indicate that inverse scope is not completely unacceptable in Chinese, and that in both languages there is a hierarchy in the ability of the three quantifiers tested to take wide scope: *every/meige* > *all/suoyoude* > numeral quantifiers. In addition, a goal or location object seems to get a wide-scope interpretation more readily than a theme object.

LY&Wb conclude that the difference between English and Chinese with regard to scope determination must lie in language-specific sensitivities to the inherent lexical properties of quantifiers and to the thematic hierarchy. They reject analyses such as Aoun & Li (1993) and Hornstein (1995), which ascribe scope differences between Chinese and English to a single syntactic parameter, and which, in addition, do not account for the different scopal properties of different quantifiers.

3.1.2.1. Comment

LY&Wb's native English results on sentences in the theme object condition with *every* as the object quantifier can be compared with those for the similar condition in Kurtzman & MacDonald (1993), namely their type (127b), *A ...every* order with an action verb. In Kurtzman & MacDonald, inverse scope was accepted around 25% of the time, while in LY&Wb, the rate of consistent acceptance of inverse scope was 44.4%. The higher rate of acceptance obtained by LY&Wb could be evidence that their precaution of using a four-point judgement scale rather than a binary scale was effective: participants who might have rejected inverse scope outright rather than pick 'yes' on a 'yes'/'no' scale may have used the intermediate judgement options to indicate that they at least found inverse scope possible. The use of a graded judgement scale thus seems useful in an investigation of whether or not inverse scope is possible.

Nonetheless, the acceptance rate of 44.4% is still rather low, considering that inverse scope is theoretically possible in English (at least with the classic universal quantifier, *every*). This result may reflect the reality that inverse scope readings are simply hard to get, even when provision is made for eliciting less obvious judgements. However, it is conceivable that the nature of the task could also have affected the results. Specifically, the task was very concentration-intensive. The participants would have had to exert considerable effort in order to read, visualise and judge the 100 interpretations of the doubly-quantified test sentences (e.g., 133A, B & C). Moreover, no contexts were provided to facilitate visualisation of the scope interpretations. LY&Wb do not report on any measures to control for fatigue or random answering. They note (LY&Wb:181, fn) that the non-inverse scope interpretations of each test sentence were considered to serve as distractors, but they do not indicate whether the results for these distractors were examined for possible indications of random answer patterns. The results do not lead us to suspect a high level of random answering (since, if there had been, no distinct patterns of answers would have emerged). Nonetheless, considering the concentration-intensive nature of the task, inclusion of a relevant control measure would have increased confidence that the rates presented were not influenced by factors such as fatigue.

Turning to LY&Wb's conclusion, their claim that their results indicate a crucial role for quantifier type and thematic role in determining scope interpretation indeed seems reasonable, given the clear effects of these variables in the results. However, it seems hasty to rule out a role for parameterisation in scope differences between Chinese and English. The native Chinese speakers' rates of consistent acceptance of inverse scope of up to 36.7% in the goal/location object condition are not necessarily problematic for the accounts of Aoun & Li (1993) or Hornstein (1995) (detailed in Chapter 2). Both accounts note that passives in Chinese are ambiguous, and show how this ambiguity can be accounted for within their frameworks. Like the grammatical subject of a passive, the grammatical subject of an unaccusative is argued, under the Unaccusative Hypothesis (Burzio 1986; Perlmutter 1978), to originate as the internal argument of the verb (i.e., in object position). Thus, under Aoun & Li (1993) and Hornstein (1995), scope ambiguity is predicted for Chinese unaccusative QP-QP sentences analogously to Chinese QP-QP passives. However, ambiguity is not predicted with action verbs, so the 26.7% consistent acceptance of inverse scope in native Chinese (representing eight out of the 30 participants) is a problem for these accounts. Aoun & Li (1993) and Hornstein (1995) assume that subjects remain in VP in Chinese, a consequence of which is that there is no way for a (thematic) object to c-command, and hence take scope over, a subject. That eight out of 30 native Chinese participants in LY&Wb's study allowed O>S scope could suggest that Aoun & Li's and Hornstein's claim is too strong. Alternatively, it could be that some other processing mechanism (perhaps along the lines of Hayashishita 1999; 2000a; 2000b; see Chapter 2) enabled those individuals to accept inverse scope at an extra-syntactic level, despite the syntactic constraints.

In summary, the results of LY&Wb, like those of Kurtzman & MacDonald (1993) suggest that scope determination is the result of more than one factor. However, before ruling out a role for proposals such as Aoun & Li (1993) and Hornstein (1995), more evidence from other studies, including the results of the present study reported in Chapters 4 and 5, should be considered.

3.1.3. Sano (2004)

The aim of Sano's (2004) study is to investigate the L1 acquisition of scope interpretation in Japanese QP-QP sentences. Specifically, he aims to discover (i) whether children have knowledge of the lack of inverse scope in the SOV QP-QP sentence in (134); and (ii) whether children know that either the subject or the object QP can take wide scope in the scrambled OSV version of (134) given in (135) (Sano 2004: 427 (15)–(16)).

134. Dareka-ga dono neko-mo tukamaeta. *Interpretation*: S>O; *O>S
someone-NOM every cat-QPt caught
'Someone caught every cat.'

135. Dono neko-mo dareka-ga tukamaeta. *Interpretation*: S>O; O>S
every cat-QPt someone-NOM caught
'Someone caught every cat. (*scrambled*)'

Twenty-two monolingual Japanese-speaking children (age 4;1–6;5) participated in the study, as well as ten adult Japanese non-linguists. The task was a truth-value judgement task. Two series of three pictures were prepared, one depicting the subject-wide interpretation of (134) and (135), the other depicting the object-wide interpretation. The first picture for each scenario is the same: three cats and four children are depicted. The faces of only three children are visible. The experimenter explains that these three children are called (for example) Taro, Hanako, and Jiro, but the name of the fourth child is unknown. In the subject-wide-scope scenario, the subsequent two pictures show that the nameless child catches all three cats. In the object-wide-scope scenario, the subsequent two pictures show that each named child catches one cat. At the end of each presentation of the series of pictures, a puppet describes what happened, using one of the sentences in (134) or (135). The participant is asked to indicate whether the puppet's description is right by feeding him candy (for a correct description) or a stone (for a wrong description).

The adult control results support the theoretical claims about scope interpretations in Japanese QP-QP sentences: all ten adults accepted the canonical SOV sentence (134) with subject-wide scope and the scrambled OSV sentence (135) with both subject-wide and object-wide scope. However, the canonical SOV sentence (134) with object-wide scope was rejected by eight of the ten adults. (The remaining two accepted object-wide scope on this sentence.) The result for the children is similar for the three sentence/scope combinations accepted by the adults: at least 15 of the 20 children accepted the canonical SOV sentence (134) with subject-wide scope and the scrambled OSV sentence (135) with both subject-wide and object-wide scope. However, only six of the 20 children rejected the canonical SOV sentence (134) with object-wide scope. In other words, 30% of the children rejected this condition, compared with 80% of the adults.

Sano (2004) suggests that the children's apparent violation of the 'ban' on object-wide scope in Japanese SOV QP-QP sentences may indicate that they have not yet acquired the lexical properties of the Japanese QP *dono N-mo* 'every N'. Specifically, he notes the difference between English *every* and *all*, whereby *every N* can readily take object-wide scope while *all the N* cannot (as noted by Ioup 1975 and as observed in the experimental data on English in the present dissertation, Chapters 4–5). He speculates that, in the child grammar, *dono ... mo* has whatever property allows English *every* to take object-wide scope, while in the adult grammar, it has whatever property prevents English *all* from taking object-wide scope.

3.1.3.1. Comment

One avenue Sano (2004) does not explore in seeking to account for the children's non-adult-like acceptance of object-wide scope in Japanese SOV QP-QP sentences is the possibility that the result comes from a 'yes' bias. In order for the children to provide an adult-like response to the object-wide scope interpretation of the sentence in (134), they would have had to tell the puppet that his description of the story was wrong. It is well known in experimental language acquisition research that children may have an aversion to giving 'no' answers, perhaps because they want to be amenable to the experimenter by saying 'yes', or perhaps because, in a task like the present one, they don't want to 'punish' the puppet by giving it a stone (see, e.g., Crain & Thornton 1998). A useful addition to Sano's experiment would, therefore, have been a distractor item that unambiguously required a 'no' response. If the children successfully rejected the puppet's description on that distractor, this would show that they were at least capable of providing a 'no' answer. Without such a control, the possibility that the child results may be due to a 'yes' bias casts doubt on their validity.

Of key interest in the present dissertation are the adult results. Although only two sentences are judged—one canonical SOV sentence (134) and one scrambled

OSV sentence (135)—the judgements were largely in accordance with the claims set out in Chapter 2 about scope interpretation in Japanese. Namely, subject-wide scope was accepted on both the SOV and the OSV sentence, while object-wide scope was rejected on the SOV sentence and accepted on the OSV sentence. As will be seen in Chapters 4 and 5, this result provides confirmation—albeit on a small scale—of the reliability of the results obtained for native Japanese in the present study.

3.2. L2 knowledge of quantifier scope interpretation

3.2.1. Lee, Yip & Wang (1999a)

Lee, Yip & Wang (1999a) (henceforth LY&Wa, in this section) use the same methodology as Lee, Yip & Wang (1999b) to investigate knowledge of quantifier scope interpretation in L1 Chinese-L2 English interlanguage. The aim of the investigation was to discover whether learners demonstrate any evidence of L1 transfer by rejecting inverse scope in English, since it is not allowed (or at least it is dispreferred, given the results of Lee, Yip & Wang 1999b) in native Chinese. In addition, the study investigated whether Chinese learners of English show the same sensitivity to the effects of quantifier type and thematic role in assigning inverse scope as native English speakers do.

The test battery was smaller than in Lee, Yip & Wang (1999b), with 76 interpretations for judgement instead of 100, due to exclusion of the test types with *every* or *all* as the subject (132d–e). The participants were 32 intermediate Chinese learners of English, 31 advanced Chinese learners of English, and a control group of 27 native speakers of English.[96] The three groups received identical instructions in English, and the whole test was presented in English. Table 4 (based on LY&Wa: 52) shows the overall rates of acceptance (i.e. selection of '1', '2' or '3' on the judgement scale, as described in Section 3.1.2) of inverse scope in the six test conditions.[97]

[96] The intermediate learners were first-year university students and the advanced learners were fourth-year university students. Mean group scores on a 30-blank cloze test verifies that the two groups are distinct: the intermediate group mean was 7.72 (SD = 2.65) and the advanced, 16.29 (SD = 3.25), the difference between these means being statistically significant ($F = 131.94$, $p < .0001$) (LY&Wa: 49). The native English group seems to be the same group reported on in Lee, Yip & Wang (1999b) since the consistent rates of acceptance reported on for the goal/location condition (the only condition for which consistent acceptance is reported) are the same as in the former study.

[97] Although LY&Wa provide 'consistent acceptance' rates for the test items in which the object had a goal or location thematic role (LY&Wa: 51), only the group means are given for the items where the object had a theme role: hence, only group means are presented in Table 4.

Table 4: Mean rates (%) of acceptance of inverse scope with different object quantifiers and object roles, in L2 English (L1 = Chinese) and native English

object QP	intermediate L2 English (n=32)		advanced L2 English (n=31)		native English (n=27)	
	object role: theme	goal/loc	***object role:*** theme	goal/loc	***object role:*** theme	goal/loc
every/meige	53.1	78.9	53.2	77.4	58.4	79.9
all/suoyoude	27.3	49.9	13.8	46.8	25.9	69.5
NumP	31.2	56.2	22.6	38.7	20.4	50.0

It is clear from Table 4 that both groups of learners demonstrated rates of acceptance of inverse scope comparable with those of the native English speakers. LY&Wa report no significant differences between the three groups (LY&Wa: 53). Thus, the results show no evidence of transfer of the low acceptance of inverse scope in Chinese to the Chinese/English interlanguage. However, LY&Wa note that this result does not rule out L1 knowledge being transferred to the interlanguage at the initial state of acquisition (as hypothesised by Schwartz & Sprouse 1996). It could be that the intermediate learners were already too advanced for any transfer effect to be detected: their interlanguage could have already undergone restructuring with respect to the differences in relative acceptability of inverse scope between English and Chinese.

Regarding the influence of quantifier type and thematic role of the object, the learners' sensitivities are very similar to those of the native English speakers. Both learner groups behaved like the native English group in allowing inverse scope significantly more with goal or location objects than with theme objects, with all quantifier types (LY&Wa: 52). In addition, the learner groups allowed *every* to take inverse scope significantly more than *all* or a numeral quantifier, with both theme and goal/location objects. Overall, the highest rates of acceptance of inverse scope are found for all three groups with the quantifier *every* for goal or location objects (78.9%, 77.4% and 79.9% acceptance by the intermediate learners, the advanced learners and the native English controls, respectively); and the next highest rates for *every* with theme objects (53.1%, 53.2% and 58.4% acceptance by the intermediate learners, the advanced learners and the native English controls, respectively). In short, the learners show native-like sensitivity to the influence of thematic role and the lexical properties of quantifier type in assigning inverse scope.

3.2.1.1. Comment

The fact that LY&Wa did not find an L1 transfer effect (i.e., the native-Chinese-speaking learners of English did not reject object-wide scope in English) may not be surprising on consideration of how the learners might acquire the availability of inverse scope in English. Firstly, the acquisition problem facing Chinese learners of English is not a poverty-of-the-stimulus problem. Since inverse scope is available in English (at least more so than in Chinese), it is not inconceivable that the learners might come across direct evidence for inverse scope in the input. Secondly, learners could acquire the availability of inverse scope indirectly. Under the proposals of Aoun & Li (1993) and Hornstein (1995), inverse scope is available when subjects raise from a VP-internal position to IP. Assuming transfer of the L1 grammar to the interlanguage at the start of L2 acquisition, the VP-internal subject account predicts that Chinese learners of English will acquire English inverse scope when evidence motivates the restructuring of their interlanguage grammar from a non-subject-raising grammar to a subject-raising grammar. Such evidence would be provided in the form of auxiliaries and modals between the subject and the verb.[98] Since auxiliaries and modals occur plentifully in English, the relevant restructuring could be motivated at an early stage of acquisition—conceivably well before the learners reach the intermediate level of LY&Wa's study. The results therefore do not provide evidence against the parameterisation accounts such as Aoun & Li's (1993) and Hornstein (1995).

As with Lee, Yip & Wang (1999b), a possible flaw in this study is that it did not include any measure to control for random answering due to fatigue or other factors. The processing burden for the learner groups is particularly heavy, since they had to read, comprehend, and judge the test sentences and the interpretations in their non-native English. The fact that LY&Wa did not present rates of consistent acceptance of inverse scope for the test items in which the object had a theme role (see footnote 97) may have been due to the levels of individual consistency not being very high. Since the mean group rates of the intermediate and advanced learners for *every* in the theme condition were around 50% (53.1% and 53.2% respectively), it would be useful to know whether these means reflect a lack of consistent responses

[98] The existence of verbal elements such as auxiliaries between the subject and the verb in English is a prime argument for VP-to-IP subject-raising in English (e.g., Kitagawa 1986). Chinese does not have such auxiliaries.

throughout each group. Such inconsistency could potentially arise from lack of concentration, which results on distractor items could have helped to identify. Nonetheless, since the mean group rates of acceptance of inverse scope were strikingly high in the goal/location condition with *every* as the quantifier (78.9%, 77.4% and 79.9% for the intermediate learners, advanced learners and native controls, respectively), all groups clearly distinguished between goal/location objects and theme objects. Thus, even if the results in the theme condition contain some randomness, this randomness did not affect all test conditions.

3.2.2. Miyamoto & Yamane (1996)

Miyamoto & Yamane (1996) (henceforth M&Y, in this section) investigate whether Japanese learners of English have native-like knowledge of the interpretation contrast between English [*wh*-subject ... QP-object ...] questions and [*wh*-object ... QP-subject...] questions. The contrast (previously outlined in Chapter 2) is exemplified as follows:

136.	a. Who has everything?	unambiguous: *wh*>QP
	b. What does everyone have?	ambiguous: *wh*>QP; QP>*wh*

The examples in (136) indicate that [*wh*-subject ... QP-object ...] questions (136a) allow only an individual reading (e.g., an answer such as *Jane does*), while [*wh*-object ... QP-subject...] questions (136b) allow both individual and pair-list readings (e.g., answers such as *Apples* or *Jane has apples, Bill has bananas, Sara has oranges*...). M&Y assume that *wh*-QP scope interactions are governed by the UG mechanisms of *wh*-movement, QR, and Aoun & Li's (1989) Scope Principle (137):

137. *The Scope Principle* (Aoun & Li 1989: 141)
A quantifier A has scope over a quantifier B in case A c-commands a member of the chain containing B.

They give the LF representations in (138) for (136a–b) respectively:[99]

138. a. [CP who$_i$ [IP t_i [VP everything$_j$ [VP has t_j]]]]
b. [CP what$_i$ [IP everyone$_j$ [IP t_j [VP have t_i]]]] (based on M&Y: 494 (2))

They explain that, by the Scope Principle (137), *who* takes scope over *everything* in (138a), but not vice versa: hence, only the individual interpretation (*wh*>QP) holds. In (138b), *what* takes scope over *everyone*, and *everyone* takes scope over the trace of *what*, thus both the individual (*wh*>QP) and pair-list (QP>*wh*) interpretations hold.

The premise of M&Y's investigation is that, if Japanese learners of English demonstrate knowledge of the interpretation contrast between [*wh*-subject ... QP-object ...] questions and [*wh*-object ... QP-subject ...] questions, this will show that they have knowledge of the UG mechanisms of *wh*-movement, QR, and the Scope Principle, and hence, that UG is available in L2 acquisition. They note, citing Hoji (1985), that, while knowledge of the unambiguity of English [*wh*-subject ... QP-object ...] questions (136a) could potentially transfer from Japanese, Japanese [*wh*-object ... QP-subject ...] questions lack a pair-list reading (as detailed in Chapter 2), hence knowledge of pair-list readings in English by native speakers of Japanese could not be ascribed to L1 transfer.

A truth value judgement task was used to investigate learners' knowledge of *wh*-QP interpretation. For each test item, a picture was shown, and two puppets (manipulated by the experimenters) held a dialogue about the picture. The dialogue included a *wh*-QP question, as exemplified in (139) for a [*wh*-subject ... QP-object ...] question (M&Y: 497 (7)):

[99] Note, however, that these representations are a simplification. Aoun & Li (1989) do not, in fact, address *wh*-QP interactions. They are addressed in Aoun & Li (1993), where the Scope Principle is modified as in (i) (previously cited in Chapter 2):

i. *The Scope Principle* (Aoun & Li 1993: 88 (50))
An operator A may have scope over an operator B iff A c-commands B or an A'-element co-indexed with B.

The modified principle in (i) does not allow for elements in A-positions, such as the object trace in (138b), to play a role in scope interpretation. See Chapter 2 for details of Aoun & Li's (1993) account of (136a–b).

139. (*Picture*: *three boys, Tom, Ken, and Neil, holding playing cards. Tom has a club, a heart and a diamond, Ken has a diamond and a heart, Neil has a club, a heart, a spade and a diamond.*)
Puppet 1: These guys are playing cards. Hey, who has everything?
Puppet 2: Let me see... Tom has a club, a heart and a diamond. Ken has a diamond and a heart. Neil has a club, a heart, a spade and a diamond. Oh, I know who has everything! Neil has everything.

Following the dialogue, the participants indicated on answer sheets whether Puppet 2's answer was *True* or *False*. Six conditions were investigated, manipulating the variables of question type ([*wh*-subject ... QP-object ...] v. [*wh*-object ... QP-subject...]), scope (*wh*>QP v. QP>*wh,* i.e., individual answer/pair-list answer), and veracity of Puppet 2's answer (true v. false). The six conditions are summarised and exemplified in Table 5 (based on M&Y: 496–498).

Table 5: Summary and exemplification of M&Y's six test types

type	question type	scope	description of picture	dialogue (abridged)		truth value
				Puppet 1	Puppet 2	
1	*Wh*-subj ... QP-obj...	*wh*>QP	Neil has four playing cards, one of each suit. Two other boys have two and three cards each.	Who has everything?	Neil has everything.	True
2	*Wh*-subj ... QP-obj...	*wh*>QP	Joe, Ron and Ken are each holding different kinds of flowers. Dick has no flowers.	Who has everything?	Joe and Dick have everything.	False
3	*Wh*-subj ... QP-obj...	*QP>*wh*	Patty has a spoon, Jane has a spoon and a fork, Sue has a knife, a fork and a spoon.	Who has everything?	Patty has a spoon, Jane has a spoon and a fork, Sue has a knife, a fork and a spoon.	False
4	*Wh*-obj ... QP-subj	*wh*>QP	Amy loves A & B. Meg loves A, B, C and D. Vicky loves B.	Who does everyone love?	Everyone loves B.	True
5	*Wh*-obj ... QP-subj	*wh*>QP	Vinny has a bowling ball and bowling shoes. Mike has a bowling ball. Richard has bowling shoes.	What does everyone have?	Everyone has a bowling ball and shoes.	False
6	*Wh*-obj ... QP-subj	QP>*wh*	Mark has a racquet and a ball. James has balls. Donald has a racquet and balls.	What does everyone have?	Mark has a racquet and a ball. James has balls. Donald has a racquet and balls.	True

The crucial test type is Type 6, since this represents the combination that cannot transfer from native L1 knowledge: [*wh*-object ... QP-subject ...] questions with a pair-list (QP>*wh*) reading. Types 2 and 5, which are both false because Puppet 2's answer does not match the picture, can potentially be used as control items: if the participants mark items of this type as *True*, it could be due to inattention or failure to understand the test procedure. Any participants answering in this way could thus be excluded from the analysis—although, as is shown below, this was not necessary, since items of these types were always answered correctly. In contrast to Types 2 and 5, Type 3—a pair-list reading of [*wh*-subject ... QP-object ...] questions—is not false because of a mismatch between the puppet's description and the picture; it is false because the reading is grammatically incompatible with the question type.

Three items for each of Types 1, 3, 4 and 6 were created, along with two for each of the control Types, 2 and 5. Nineteen filler items were randomly mixed with the 16 test items, in order to disguise the focus of the experiment. The participants were 15 adult learners of English whose native language was Japanese. All were students registered on academic courses at Ohio University. They had had six years of English instruction in Japan, from age 12, and their TOEFL scores ranged from 450–600. In addition, five native English speakers made up a control group.

The rates of correct responses (i.e., responses matching those given in the 'truth value' column of Table 5) are provided in Table 6 (based on M&Y: 499).

Table 6: Correct response rates (raw numbers of correct responses in brackets)[a]

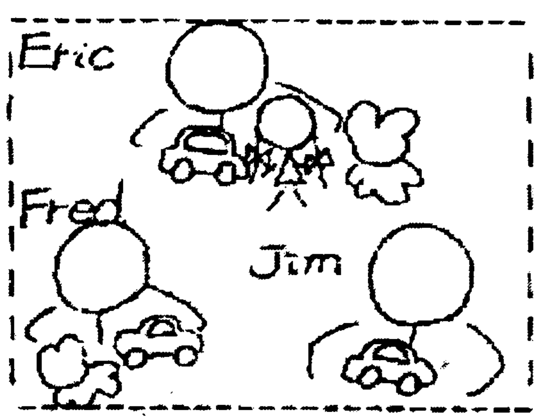
Participants	Type 1 *Wh*-subj, *wh*>QP, T (n=3)	Type 2 *Wh*-subj, *wh*>QP, F (n=2)	Type 3 *Wh*-subj, *QP>*wh*, F (n=3)	Type 4 *Wh*-obj, *wh*>QP, T (n=3)	Type 5 *Wh*-obj, *wh*>QP, F (n=2)	Type 6 *Wh*-obj, QP>*wh*, T (n=3)
Japanese learners of English (n=15)	100.00% (45/45)	100.00% (30/30)	95.56% (43/45)	97.78% (44/45)	100.00% (30/30)	71.11% (32/45)
Native speakers of English (n=5)	100.00% (15/15)	100.00% (10/10)	86.67% (13/15)	100.00% (25/25)	100.00% (10/10)	93.33% (14/15)

[a] '*Wh*-subj' = [*wh*-subject ... QP-object...] question, '*Wh*-obj' = [*wh*-object ... QP subject...] question, 'T' = 'True', 'F' = 'False'.

The results in Table 6 show that high levels (>70%) of correct responses were demonstrated by both the learners and the native controls on all test types. However, the greatest difference between learners and native controls occurs on Type 6—the [*wh*-object ... QP-subject ...] question with a pair-list answer. On this type, the learner rate of correct responses was 71.11% compared with 93.33% by the native English group: a difference of over 22%, compared with differences of less than 9%

on the other test types. M&Y argue that the learners' depressed rate of correct responses on this type was due to an experimental effect, since 11 out of the 13 incorrect responses occurred on a single test item. The one incorrect native response also occurred on that item. This item is illustrated in (140) (M&Y: 505 (iv)):

140.



Puppet 1: Nice toys! Cars, a doll, and teddy bears. Hey, what does everyone have?

Puppet 2: Let me see... Oh, I know what everyone has! Eric has a car, a doll and a teddy bear, Fred has a teddy bear and a car. Jim has a car!

M&Y point out that the two preceding [*wh*-object ... QP-subject ...] questions (items 17 and 22 on the test) had had individual answers (i.e., the *wh*>QP scope condition, rather than the QP>*wh* scope condition of (140)). They suggest that this may have created an expectation of the next [*wh*-object ... QP-subject ...] question (item 32) also having an individual interpretation, thus giving rise to the large number of *False* responses for this item (M&Y: 501).[100] M&Y conclude that, since, overall, the

[100] However, doubt is cast on this explanation by the fact that nine other test items intervened between this item (item 32) and the preceding [*wh*-object ... QP-subject ...] question (item 22). It seems unlikely that an expectation about a particular question type could be maintained over so many intervening items. Also, an effect like this would be expected in the control group as well, yet only one of the five native English speakers judged the item incorrectly. Given this learner/native difference on this test item, it might be supposed that L1 transfer played a role: since pair-list readings are not readily available in Japanese [*wh*-object ... QP-subject ...] questions, the notion that the learners were influenced by their L1 cannot be dismissed. However, this analysis is also flawed in that L1 influence should manifest itself in all three Type 6 items, if it is the cause.

A third option is that the learners misunderstood the picture, and judged the puppet's answer false because they felt it did not match the picture. Looking critically at the picture, the learners may have identified the teddy bear held by Fred as some other animal than a bear—although M&Y clearly designed the puppet's dialogue to avoid such misidentification, by including a statement of what the picture represents (*Puppet 1: Cars, a doll, and teddy bears!*). If misunderstanding of the picture is the reason for the learners' high rate of incorrect answers, the question again arises of why this did not happen with the English control group. Only a speculative solution can be offered: perhaps, since pair-list readings are in fact the most obvious readings of [*wh*-object ... QP-subject ...] questions in English (see empirical support of this claim in Chapter 5), the native English speakers were more readily able to suspend disbelief and accept that the picture depicts what Puppet 1 claims it depicts, since that leads to the preferred reading of this question type; however, for Japanese learners of English, pair-list readings may be less preferred due to L1 influence, and, in addition, the very act of taking a test in English may provoke extra attention to detail in an attempt to get the answers right. This combination may have predisposed the learners to be highly attentive to possible mismatches between the picture and the dialogue.

The best solution for item 32 may have been to exclude it from the analysis. Without this item, the learners' rate of correct answers to Type 6 questions rises to 93.33% (28 out of 30).

Japanese learners of English demonstrate native-like knowledge of scope interpretation in *wh*-QP questions, *wh*-movement, QR and the Scope Principle must be operative in the interlanguage, and therefore UG must be available in L2 acquisition.

3.2.2.1. Comment

A criticism of M&Y's study is that it comes up against some of the problems set out in Schwartz & Sprouse (2000) (outlined in Chapter 1). Schwartz & Sprouse (2000: 197) point out that 'a UG-compatible analysis of an L2 data set does not, on its own, constitute an argument for the Strong UG hypothesis [i.e., that L2 acquisition is constrained by UG]'. Only when the phenomenon under investigation can be demonstrated to be an L2 poverty-of-the-stimulus problem does acquisition of that phenomenon implicate the involvement of UG, since, logically, nothing else in the environment could account for the relevant acquisition. In M&Y's study, acquisition of the pair-list interpretation of English [*wh*-object ... QP-subject ...] questions by native speakers of Japanese is not a poverty-of-the-stimulus problem. Although the facts are underdetermined by the L1 (since Japanese [*wh*-object ... QP-subject ...] questions do not have a pair-list reading), they are not underdetermined by the input, which might conceivably contain [*wh*-object ... QP-subject ...] questions in clear pair-list contexts. Given the potential existence of such evidence, the fact that the learners demonstrated knowledge of pair-list interpretations in English does not prove that UG constrains L2 acquisition: an alternative account could be that they induced the availability of pair-list interpretations from the input.

This problem seriously undermines M&Y's conclusion. The study is nonetheless interesting from the point of view of its experiment design. The use of puppets and pictures to determine the scope reading of the *wh*-QP questions avoids the potentially burdensome task of reading written scope interpretations. In addition, the task seems well designed to identify possible non-linguistic answering strategies. One possible strategy would have been to ignore the question and answer, but to select *True* when Puppet 2 correctly described the picture, *False* when the puppet's description did not match the picture. However, such a strategy would not have worked on Type 3 ([*wh*-subject ... QP-object ...] questions with a grammatically incompatible pair-list answer). On these items, Puppet 2's description matched the picture, but the correct answer was still *False*, due to the incompatibility of a pair-list reading with this question type. If the learners had been using a strategy of choosing *False* only when there was a mismatch between the puppet's description and the picture, they would have selected *True* on these items. Their 95.56% rate of selection of *False* shows that they did not adopt such a strategy. In addition, the test included items to control for inattention (Types 2 and 5), as already described. Thus M&Y's test design included control measures that serve to increase confidence that the test did indeed measure what it set out to measure.

3.2.3. Miyamoto & Takata (1998)

Miyamoto & Takata (1998) (henceforth M&T, in this section) also investigate the interpretation of English *wh*-QP questions by Japanese learners of English. They aim to test the hypothesis that elementary-level Japanese learners of English will not allow pair-list readings in English [*wh*-object ... QP-subject ...] questions, unlike the intermediate/advanced learners investigated in Miyamoto & Yamane (1996) (M&T: 514). This hypothesis is based on findings from Martohardjono & Gair (1993) and White (1992) which suggest that when L2 learners whose L1 lacks overt *wh*-movement (such as Japanese) acquire a language with overt *wh*-movement (such as English), they initially produce *wh*-phrases directly in Spec,CP and fill the position from which the *wh*-phrase should have moved with a resumptive *pro*. If this is correct, there will be no *wh*-traces in the *wh*-QP questions of early L1 Japanese-L2 English interlanguage, as illustrated in the representations in (141) (M&T: 514 (12)–(13)).

141. a. *[wh-obj ... QP-subj ...]:* [$_{CP}$ Who$_1$ did [$_{IP}$ everyone$_2$ meet *pro*$_1$]]
     b. *[wh-subj ... QP-obj ...]:* [$_{CP}$ Who$_1$ [$_{IP}$ *pro*$_1$ met everyone$_2$]]

Recall that the pair-list reading in [*wh*-object ... QP-subject ...] questions in native English is argued to be available due to the QP-subject c-commanding a trace of the *wh*-object.[101] In the representation in (141a), since there is no *wh*-trace, the pair-list reading is predicted to be unavailable.

[101] M&T adopt Chierchia's (1991, 1993) account (already detailed in Chapter 2), in which the *wh*-trace contains an argument (*pro*) which must be c-commanded and bound by a QP in order for a pair-list reading to obtain. Whether this account is adopted, or the simplified account based on Aoun & Li (1989) used by Miyamoto & Yamane (1996), the crucial point is that pair-list readings are claimed to require c-command of a *wh*-trace by the QP.

M&T investigate *wh*-QP questions with psych-verbs (e.g., *worry*, *please*), as well as with non-psych-verbs. With psych-verbs, a pair-list reading is possible in [*wh*-subject … QP-object …] questions, as well as in [*wh*-object … QP-subject …] questions (Chierchia 1993; Kim & Larson 1989). This is illustrated in (142) (QP>*wh* indicates the availability of the pair-list reading):

142. a. *[wh-obj ... QP-subj ...]*: Who did everyone worry? *wh*>QP; QP>*wh*
b. *[wh-subj ... QP-obj ...]*: Who worries everyone? *wh*>QP; QP>*wh*

Psych-verbs are argued (Belletti & Rizzi 1988) to be a type of unaccusative, in which the grammatical subject originates as an internal argument of the verb. This results in the structures in (143a–b) for (142a–b).

143. a. [CP Who$_1$ did [IP everyone$_2$ [VP [V' worry t_2] t_1]]]
b. [CP Who$_1$ [IP t_1 [VP [V' worry t_1] everyone$_2$]]]

In (143a), *who* c-commands *everyone*, and *everyone* c-commands a trace of *who*, so the pair-list reading is available just as in [*wh*-object ... QP-subject ...] questions with non-psych-verbs. In (143b), due to the subject originating as the sister of the verb, the QP-object *everyone* c-commands the *wh*-subject-trace thus allowing QP>*wh* scope, and hence pair-list interpretations, unlike in [*wh*-subject ... QP-object ...] questions with non-psych verbs. However, if the interlanguage grammar lacks *wh*-movement, and hence also *wh*-traces, pair-list readings are predicted to be absent with psych-verb *wh*-QP questions, just as described for [*wh*-object ... QP-subject ...] questions with non-psych-verbs (141a–b).

M&T investigate learners' interpretations of the four question types in (144):

144.

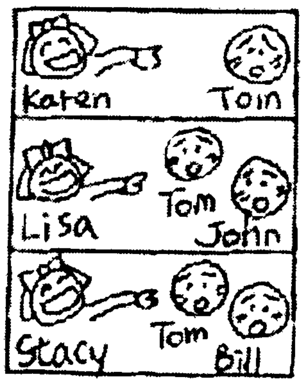
type	**verb type**	**question type**	**example** (M&T: 515–6)
1	non-psych-verb	*Wh*-obj … QP-subj…	Who did everyone laugh at at his party?
2	non-psych-verb	*Wh*-subj … QP-obj…	Who hit everyone at his party?
3	psych-verb	*Wh*-obj … QP-subj…	Who did everyone surprise at his party?
4	psych-verb	*Wh*-subj … QP-obj…	Who worried everyone at his party?

All but Type 2 in (144) are ambiguous, allowing both individual and pair-list answers. Type 2 allows only individual answers. Fourteen items were created for each of Types 1 and 2, and three items for each of Types 3 and 4. Twenty-four filler items were randomly mixed with the test sentences, creating a 60-item test. Each stimulus question was presented aurally and was accompanied by a picture. Participants indicated their interpretation of the question by selecting an answer from four multiple-choice questions. These are illustrated for a Type 1 test item in (145) (M&T: 515 (17)):

145.

Who did everyone laugh at at his party?

a. Tom.
b. Karen laughed at Tom, Lisa laughed at John, Lisa laughed at Tom, Stacy laughed at Bill, and Stacy laughed at Tom.
c. Either (a) or (b).
d. None of the above.



As (145) shows, the multiple choice questions offered the participant the choice of selecting only an individual answer (a), only a pair-list answer (b), both an individual and a pair-list answer (c), or neither an individual nor a pair-list answer (d).

The participants were 15 Japanese students at Ohio University enrolled either in intensive ESL courses or in academic courses. Their TOEFL scores ranged from 400–547.[102] Three native English speakers acted as a control group.

M&T report that the three native English speakers correctly observed the ambiguity (Types 1, 3, and 4) and lack of ambiguity (Type 2) of the four question types (M&T: 516). The learners, on the other hand, tended to fail to allow pair-list answers to the ambiguous questions. The learners' rates of selection of each response option (a)–(d) for the four question types are shown in Table 7 (based on M&T: 516 (21)). Following M&T, Table 7 divides the learners into two proficiency groups: Group JE1 comprises 11 learners whose TOEFL scores range from 400 to 490; Group JE2 comprises four learners whose TOEFL scores range from 520 to 547.

[102] M&T assert (M&T: 511) that the L2 participants in this study are 'elementary' learners, while those in Miyamoto & Yamane (1996) were 'intermediate or advanced'. It is not clear how this claim can be justified, given that the range in TOEFL scores of the two groups are very similar: 400–547 in M&T, 450–600 in Miyamoto & Yamane (1996). M&T do not state whether there was any significant difference between the TOEFL scores of the learner groups in the two studies; nor are group mean TOEFL scores reported. This is a problem for M&T's analysis. It is discussed further in Section 3.2.3.1.

Table 7: Rates (%) of selection of the multiple-choice responses (a)–(d) by Japanese learners of English[a]

		response types			
		(a) individual answer only	**(b)** pair-list answer only	**(c)** both (a) and (b)	**(d)** none of (a), (b) or (c)
JE1 (n=11)	**Type 1** non-psych verb, *wh*-obj	57.69	13.85	7.69	20.77
	Type 2 non-psych verb, *wh*-subj	85.43	4.64	6.62	3.31
	Type 3 psych verb, *wh*-obj	48.49	15.15	9.09	27.27
	Type 4 psych verb, *wh*-subj	79.74	6.98	6.98	9.30
JE2 (n=4)	**Type 1** non-psych verb, *wh*-obj	67.93	11.32	0	20.75
	Type 2 non-psych verb, *wh*-subj	85.71	8.93	3.57	1.79
	Type 3 psych verb, *wh*-obj	37.50	18.75	0	43.75
	Type 4 psych verb, *wh*-subj	87.50	0	0	12.50

[a] '*wh*-obj' = [*wh*-object ... QP subject...] question; '*wh*-subj' = [*wh*-subject ... QP-object...] question

Table 7 shows that both learner groups demonstrated low rates (<19%) of selection of options (b) and (c)—the two response options which showed acceptance of pair-list answers—on all four test types. Rates of selection of (a)—an individual answer—were particularly high (>79%) in both groups for [*wh*-subject ... QP-object ...] questions, Types 2 and 4. The [*wh*-object ... QP-subject ...] questions, Types 1 and 3, yielded more variation in response type, although the rate of selection of individual answers (a) was higher than the rates of selection of other response, except in the JE2 group, where option (d)—rejection of both individual and pair-list answers—was highest, at 43.75%.

M&T claim that these results confirm their hypothesis: lower proficiency Japanese learners of English tend not to allow pair-list readings of *wh*-QP questions. They conclude that this supports the theory that *wh*-movement is lacking in the early interlanguage grammar of L2 learners whose L1 does not have overt *wh*-movement.

3.2.3.1. Comment

A serious problem with M&T's analysis concerns their assumption that the learners in this study are at an 'elementary' level of proficiency while those in Miyamoto &

Yamane (1996) were at 'intermediate' or 'advanced' levels. This assumption appears to be based on TOEFL scores, yet, as already observed in footnote 102, there is considerable overlap between the TOEFL scores of the 15 learners in each study: 400–547 in M&T, 450–600 in Miyamoto & Yamane (1996). Given this overlap, it is not at all clear whether the learners in M&T were really at a distinctly different level of proficiency to those in Miyamoto & Yamane (1996). Indeed, M&T's JE2 group have TOEFL scores (520–547) that fall entirely within the range of those in Miyamoto & Yamane (1996). The fact that these four learners accepted pair-list readings of [*wh*-object ... QP-subject ...] questions with non-psych-verbs (Type 1) only 11.32% of the time appears to undermine the results of Miyamoto & Yamane (1996), where pair-list readings were accepted over 70% of the time.

I suggest that the discrepancy between the low rates of acceptance of pair-list readings by the JE2 group in M&T and the high rates by seemingly comparable learners in Miyamoto & Yamane (1996) could be a result of M&T's experiment design. M&T's test stimuli included a preposition phrase containing a possessive pronoun (*his*), as shown in the Type 1 question in (144), repeated in (146):

146. Who did everyone laugh at [at his party]?

The singular possessive pronoun and singular noun seem to induce a bias towards an individual interpretation of the question: there is one male who had one party and everyone laughed at him. The pair-list interpretation, *Karen laughed at Tom, Lisa laughed at John, ...* (see (145b)), seems unexpected, given the context of *at his party*, because it entails that *his* refers to more than one male and (presumably) that there was more than one party. While this is certainly a valid interpretation of *his party* in (146), it is not an obvious interpretation. Inclusion of the PP thus adds an extra interpretation problem to the test: one which is most easily resolved by assuming that *his* refers to a single male and that therefore the individual answer is the, possibly only, correct answer to the *wh*-QP question.[103] The high rates of selection of only an

---

[103] In fact, the use of *his* potentially entails yet another interpretation problem, since it could have either *who* or *everyone* as its antecedent. M&T's pictures seem to have been constructed so that the person or people corresponding to *who* (i.e. the object(s) of *laugh at* in (146)) are always male, while the person or people corresponding to *everyone* are always female. Masculine, singular personal pronouns are normally only used to refer to anyone other than a single male when the gender of the referent is not known, so the depiction of *everyone* as female should have avoided misinterpretation of *everyone* as the antecedent of *his*.

individual answer (multiple choice option (a)) in M&T's study could be a reflection of the fact that the PP in the test stimuli created a bias towards that answer.

This analysis does not, of course, rule out the possibility that *wh*-movement could be lacking in the early interlanguage grammar of Japanese learners of English. A reworking of M&T's experiment, omitting the PPs and using a participant group whose level was demonstrably lower than the learner group in Miyamoto & Yamane (1996), could contribute towards shedding light on the issue.

3.2.4. Dekydtspotter, Sprouse, Swanson & Thyre (1999), Dekydtspotter, Sprouse & Swanson (2001)

The study reported in Dekydtspotter *et al* (1999) and Dekydtspotter, Sprouse & Swanson (2001) (henceforth DSS&T/DS&S in this section) differs from the three preceding L2 acquisition studies in that the scope interpretation phenomenon investigated represents an L2 poverty-of-the-stimulus problem. The study thus follows the logic of Schwartz & Sprouse (2000) and aims to discover whether L2 learners can acquire knowledge of the target language facts despite poverty of the stimulus. If the relevant L2 knowledge arises, 'it must be the case that this L2 knowledge lies within the same cognitive domain as native language knowledge' (DSS&T: 162).

The phenomenon under investigation is the interpretation of French [*wh*-object ... QP-subject ...] questions in which the *wh*-object is *combien de N* 'how many N' and the subject is an NP quantified by *tous* 'all'.[104] These questions have two forms: *combien de N* can undergo *wh*-movement to the front of the sentence as a unit ('continuous *combien* extraction') (147); or, *combien* 'how many' can move alone, while *de N* remains in situ ('discontinuous *combien* extraction') (148) (DSS&T: 163 (1–2)).

147. <u>Combien de livres</u> est-ce que tous les étudiants lisent?
how many of books do all the students read
'How many books do all the students read?'

148. <u>Combien</u> est-ce que tous les étudiants lisent <u>de livres</u>?
how many do all the students read of books
'How many books do all the students read?'

---

[104] In the examples in (147–148), the quantifier *tous* 'all' appears directly before the NP, but in the actual test items (exemplified in (150–151), *tous* is a floating quantifier, separated from the subject NP by an auxiliary verb.

Whether or not the *combien* constituent occurs continuously or discontinuously affects the possible scope interpretations. With the continuous *combien* extraction (147), either S>O scope ('for each student, how many books did that student read?') or O>S scope ('for how many books did each student read those books?') is possible; with the discontinuous *combien* extraction (148), only S>O scope is possible. Thus, assuming a situation in which Student A reads Books X, Y and Z, and Student B reads Books X, Y and W, the answer to the question in (147) can be either 'Three' (i.e., each student reads three books: S>O scope) or 'Two' (i.e., Books X and Y: O>S scope), while question (148) can only be answered with 'Three'.

For native speakers of English, acquisition of the interpretation distinction between continuous and discontinuous *combien* extractions is a poverty-of-the-stimulus problem. While the English equivalent of (147) is like the French in allowing both the S>O and the O>S interpretations, English does not have a structural equivalent of the discontinuous *combien* extraction; so knowledge that discontinuous *combien* extractions lack an object-wide interpretation cannot derive from the L1. Classroom instruction does not provide evidence for this lack, as discontinuous *combien* constructions are not covered in French language teaching materials (DSS&T: 164). DSS&T (165) suggest that learners will have no reason not to assume that the discontinuous *combien* structure is a rewrite variant of the familiar continuous *combien* structure and hence allows both subject-wide and object-wide readings. Under this assumption, the target language input will never provide evidence for the lack of object-wide scope in discontinuous *combien* constructions, since failure to encounter an object-wide scope reading does not rule out such an interpretation ever being available. Thus, logically, there is no way for English learners of French to deduce that discontinuous *combien* constructions such as (148) lack an object-wide scope interpretation.

DSS&T/DS&S use a truth-value judgement task to investigate whether learners can nonetheless acquire the interpretation distinction between continuous and discontinuous *combien* extractions. Participants were presented with scenarios written in English, followed by a *combien* question and answer in French. They were asked to judge whether the answer to the question was possible by selecting 'yes', 'no' or 'cannot decide'. Two two-level variables—*combien* constituent (continuous

v. discontinuous) and scope (subject-wide v.object-wide)—were manipulated to create four test types, as listed in (149).

149. Type 1: Continuous *combien* question, subject-wide scope answer.
Type 2: Continuous *combien* question, object-wide scope answer.
Type 3: Discontinuous *combien* question, subject-wide scope answer.
Type 4: *Discontinuous *combien* question, object-wide scope answer.

Type 4 is the crucial item, since it represents the unacceptable combination of a discontinuous *combien* construction with object-wide scope. Examples of Type 3 and Type 4 test items are given in (150) and (151) (DSS&T: 166–7). (English glosses of the French questions are provided here for convenience, but they were not included in the test items.)

150. *Type 3: Discontinuous combien question, subject-wide scope answer*
It was Rachel's 12th birthday party, and the highlight was a magician who could read minds. The magician's assistant brought out a dog, a cat, a mouse, and a bird. The magician asked Rachel, Guy, and Annabelle to think of which animals they would like to have. Rachel thought of the cat and the dog. Guy thought of the bird and the dog. Annabelle thought of the mouse and the dog. The assistant then asked the magician:
"Combien est-ce que ces enfants ont tous voulu d'animaux?" (*How many animals did these children all want*?)
The magician answered:
"Deux, bien sûr." (*Two, of course.*)

151. *Type 4: *Discontinuous combien question, object-wide scope answer*
Philippe and Monique Dupont hired a magician and her assistant to entertain at their son Jean's 10th birthday party. The magician announced that she could read minds. Her assistant brought out a white rabbit, a black rabbit, a bird, a cat, and a dog. From experience, the magician and her assistant knew that the children would include the 2 rabbits in their choices. The magician then asked Jean, Marie, and Paul to think of the animals they would like to have. As predicted, Jean thought of the 2 rabbits and the bird. Marie thought of the 2 rabbits and the dog. Paul thought of the 2 rabbits and the cat. Winking, the assistant then asked the magician:
"Combien est-ce que ces enfants ont tous voulu d'animaux?" (*How many animals did these children all want*?)
Sure that the expected pattern had arisen, the magician answered:
"Deux, bien sûr." (*Two, of course.*)

The scenarios in (150) and (151) were also used for test items corresponding to Types 1 and 2 respectively, but with a continuous, instead of a discontinuous, *combien* question. Six further quadruples were created, yielding a total of 28 test items, to which 14 filler items were added. The majority of the fillers were designed to elicit 'no' responses (i.e., responses indicating that the answer in the test item was not correct), as a counterbalance to the actual test stimuli being expected to yield 'yes' responses on three of the four Types (DSS&T: 168).

The participants were 71 beginning-intermediate learners of French and 32 advanced learners of French, all with English as their L1. The beginning-intermediate group participants were all enrolled in French language courses at a US university. The advanced group participants had all had the equivalent of four years of French language instruction and had spent at least one year in a French-speaking country. In addition, 31 native speakers of French (students at a university in France) acted as a control group.

Analysis of the results showed that a number of participants, particularly in the learner groups, demonstrated high rates of rejection of the (acceptable) subject-wide scope construal of discontinuous *combien* questions (Type 3). Since this construal should be readily available, those participants who rejected it over 50% of the time were considered to lack the ability to interpret discontinuous *combien* questions at all, and their results were excluded from the final analysis. This left 21 participants in the beginning-intermediate learner group, 20 in the advanced learner group, and 30 in the native French group. These participants' rates of acceptance of the answers for each question type are presented in Table 8 (based on DSS&T: 169):

Table 8: Rates (%) of acceptance (selection of 'yes') on the four test types

Test item type	beg-int L2 French (n=21)	advanced L2 French (n=20)	native French (n=30)
Type 1 Continuous. *combien*, S>O	59.86	72.85	75.71
Type 2 Continuous. *combien*, O>S	42.18	48.57	26.19
Type 3 Discontinuous *combien*, S>O	59.18	69.28	86.67
Type 4 *Discontinuous *combien*, O>S	42.18	25.00	8.10

Of key interest is whether or not the learners differentiate between object-wide scope in the continuous *combien* condition (Type 2) and object-wide scope in the discontinuous *combien* condition (Type 4). DSS&T/DS&S report that the advanced learners, like the native French controls, do make the distinction: both groups accept object-wide scope significantly more on Type 2 items than Type 4 items (26.19% vs. 8.10%, $t(29) = 2.68$, $p = .012$ for the native French controls; and 48.57% vs. 25.00%, $t(19) = 4.28$, $p < .0005$ for the advanced learners) (DSS&T: 168–9). By contrast, the beginning-intermediate L2 group does not differentiate between the continuous and discontinuous conditions with respect to object-wide scope, demonstrating an acceptance rate of 42.18% for both Types 2 and 4. In short, a developmental pattern is observed: the less advanced learners treat discontinuous *combien* questions as if they had the same interpretation possibilities as continuous *combien/how many* questions in French or English, while the advanced learners appear to have knowledge of the lack of object-wide scope in discontinuous *combien* questions.

As described above, acquisition of the lack of object-wide scope in French discontinuous *combien* questions represents a poverty-of-the-stimulus problem for learners whose L1 is English. DSS&T/DS&S conclude that their evidence that the relevant knowledge can nonetheless be acquired by the higher proficiency learners supports the hypothesis that L2 acquisition is constrained in the same way as native language acquisition.

3.2.4.1. Comment

DSS&T/DS&S's conclusion—that, in effect, L2 acquisition is constrained by UG—is the same as the conclusion drawn by Miyamoto & Yamane (1996). However, unlike Miyamoto & Yamane (1996), DSS&T/DS&S provide evidence of native-like L2 knowledge arising under poverty of the stimulus. Consequently, their results provide a much stronger argument for L2 acquisition being constrained by UG, since there is no other way in which the relevant knowledge could feasibly be acquired. The study thus furnishes an example of the utility in L2 acquisition research of investigation of a poverty-of-the-stimulus problem: such investigations allow much clearer conclusions to be drawn about the role of UG in L2 acquisition.

The remaining comments in this section focus on aspects of DSS&T/DS&S's experimental design. As in Kurtzman & MacDonald's (1993) study, the respondents could select 'yes' or 'no' to indicate their judgement of whether or not they found each scope interpretation possible. DSS&T/DS&S also offered a 'cannot decide' option. However, unlike in Lee, Yip & Wang's (1999a, 1999b) studies in which a four-point judgement scale was used, there was no opportunity for DSS&T/DS&S's respondents to indicate intermediate levels of acceptability. It is interesting to note, therefore, that DSS&T/DS&S's native French acceptance rate for object-wide scope (inverse scope) on the continuous *combien* questions (Type 2) was rather low, at 26.19% (even though object-wide scope is theoretically acceptable), just as Kurtzman & MacDonald's (1993) native English acceptance of object-wide scope in active QP-QP sentences was also low, at only approximately 20–35% (again, even though object-wide scope is possible). By contrast, the native English speakers in Lee, Yip & Wang (1999a) demonstrated a somewhat higher overall rate of acceptance (including responses of 'possible but rare' and 'sometimes possible', as well as 'highly possible') of object-wide scope on QP-QP sentences (agent subject = NumP, theme object = *every N*), at 58.4% (See Table 4). DSS&T/DS&S's native French results thus provide provide further support for the suggestion (Section 3.1.2.1) that use of a graded judgement scale is indeed more effective than a categorical yes-no scale in eliciting judgements about the *possibility* of an interpretation and not only whether or not that interpretation is preferred. Specifically, as Lee, Yip & Wang (1999b) suggest, forcing the respondents to select either 'yes' or 'no' may lead to less obvious, but nonetheless possible, scope interpretations falling into the 'no' (or 'cannot decide') category and thus depressing rates of acceptance of theoretically possible interpretations.

A further interesting aspect of DSS&T/DS&S's experimental design is that they used the learners' native language, English, to present the scenarios for the test items. This was to ensure that the scenarios could be understood even by beginning-level learners with as little difficulty as possible (DSS&T: 165). This is indeed an important consideration, since the scenarios are crucial in providing plausible contexts for both narrow- and wide-scope construal of the object. However, a drawback to requiring the learners to read contexts in their native language is that it may encourage them to 'think in English' and perhaps use a translation strategy in addressing the French questions following each scenario—which would clearly raise the rate of acceptance of (non-target-like) object-wide scope on the discontinuous *combien* questions. In addition, since the French control group also read the scenarios in English, the possibility of this group making English-related misunderstandings cannot be ruled out (although the French control group were all taking advanced degrees in English, so this possibility is minimised). Despite these concerns, since DSS&T/DS&S's results did not yield unexpected or anomalous patterns, it seems that the main advantage of the use of English scenarios—ease of comprehension by the lower proficiency learners of French—outweighed the possible drawbacks. However, use of a design like this becomes more problematic when the number of L1s in the experiment increases, as in the present dissertation which investigates learners of Japanese whose L1s are English, Korean, or Chinese. Adopting an experimental design that uses written scenarios presented in the L1 would require that the scenarios be translated into these three languages (and probably also into Japanese for the native Japanese control groups, since, unlike in DSS&T, the existence of three L1s among the learners means that there would be no reason for using English as the presentation language for the control group). This would introduce the considerable problem of controlling for equivalence across the translated scenarios. As will be detailed in the following chapter, an experimental design using scenarios was piloted for the present research, but eventually rejected in favour of a methodology using pictures, which remain constant whatever the L1 of the participants. Controlling for translation equivalence was one reason for rejecting the scenario method.

3.3. Conclusion

Seven investigations of QP-QP or *wh*-QP scope interpretation have been presented in this chapter, three focussing on native intuitions, and four investigating scope interpretation in L2 acquisition. Discussion has focused on the findings with respect to the availability of object-wide (inverse) scope and pair-list readings, and also on aspects of the design of the experiments. The key points are summarised as follows.

The investigations of QP-QP scope interpretation in native English by Kurtzman & MacDonald (1993) and Lee, Yip & Wang (1999b) showed that object-wide scope is less readily available than subject-wide scope, despite the fact that object-wide scope is theoretically acceptable in English. In addition, Kurtzman & MacDonald's (1993) and Lee, Yip & Wang's (1999b) studies provided evidence that factors such as thematic role, argument type, and quantifier type affect the ability of a QP to take wide scope, indicating that a number of determinants are involved in scope interpretation.

Lee, Yip & Wang (1999b) also investigated native Chinese intuitions about QP-QP interactions and found that inverse scope was accepted up to 26.7% of the time (on agent-verb-theme sentences), despite many claims in the literature that Chinese lacks inverse scope. This rate is nonetheless considerably lower than the native English rate of acceptance of inverse scope on equivalent sentence types (44.4%); thus the claim that English and Chinese differ, at least qualitatively, with respect to the availability of inverse scope is upheld. It was suggested that the 26.7% rate of acceptance of inverse scope in Chinese could indicate that inverse scope contruals can be constructed at an extra-syntactic level even when the syntax rules out the relevant interpretation.

Sano's (2004) finding about adult native Japanese interpretation of QP-QP sentences was in accordance with the claims of the theoretical literature (see Chapter 2): subject-wide scope was acceptable on both the canonical SOV sentence and the scrambled OSV sentence investigated; however, object-wide scope was acceptable only on the OSV sentences, and it was rejected on the SOV sentence.

With regard to scope interpretation in L2 acquisition, Lee, Yip & Wang (1999a), Miyamoto & Yamane (1996) and Dekydtspotter *et al* (1999)/Dekydtspotter *et al* (2001) all found that learners were able (at least, eventually) to achieve native-like scope interpretation knowledge. However, only Dekydtspotter *et al* (1999)/Dekydtspotter *et al* (2001) investigated an L2 poverty-of-the-stimulus problem; thus only their findings can be argued—along the lines of Schwartz & Sprouse (2000)—to strongly implicate UG in L2 acquisition.

Miyamoto & Takata (1998) claimed that their findings demonstrated a lack of knowledge of pair-list readings in English [*wh*-object ... QP-subject ...] questions by (purportedly) elementary-level Japanese learners of English. However, some serious problems with the constructions used in the test stimuli and with the claims about the proficiency level of the learners cast doubt on the findings of this study.

All the investigations used some kind of judgement task to find out about the acceptability of different scope readings. It was noted that those studies that used a binary, yes-no judgement scale (Kurtzman & MacDonald 1993, Dekydtspotter *et al* 1999) obtained lower rates of native acceptance of (theoretically possible) object-wide scope than Lee, Yip & Wang (1999a, 1999b), who used a four-point scale and thus allowed for judgements of intermediate acceptability as well as categorical acceptable/unacceptable judgements. This was taken to support Lee, Yip & Wang's (1999b: 181) suggestion that a binary scale may lead participants to reject readings which are marked but nonetheless possible. A graded judgement scale thus seems more effective in facilitating judgments that admit the possibility of, not just the preference for, particular scope interpretations.

Lee, Yip & Wang's (1999a, 1999b) studies and Miyamoto & Yamane's (1996) study were compared with regard to the processing burden placed on the participants. The experiment in the former two studies was very long, and required intensive reading and puzzling out of written scope interpretations, potentially increasing the risk of participants losing concentration and answering randomly. In addition, the test battery did not include fillers designed specifically to control for effects such as fatigue. Miyamoto & Yamane's (1996) design, on the other hand, was shorter and used puppets and pictures to provide the contexts for the sentences to be judged, thus placing a lower processing burden on the participants. Moreover, they incorporated fillers that could be used to control for test-taking strategies or fatigue effects. The shorter, less burdensome design that also incorporates control items is clearly preferable for increasing the validity of the results.

A final design issue discussed was the use, by Dekydtspotter *et al* (1999)/Dekydtspotter *et al* (2001), of the learners' L1 (English) for presenting scenarios which provided the context for the French test items. The key advantage of this approach is that it should maximise learner understanding of the scenarios, while the disadvantage is that it may encourage the learners to think in their L1 and consequently apply translation strategies to their interpretation of the target language test stimuli following the scenario. It was also noted that presentation in the learners' L1 of context-providing scenarios may not be suitable for studies (like that of the present dissertation) in which more than one L1 is involved. Where the participants have different L1s, the scenarios will have to be translated into each L1, introducing a translation equivalence variable, which is not necessarily easy to control.

The findings of the studies outlined in this chapter, as well as the relevant design issues, are used to inform the experimental design of the investigation methodology used in the present dissertation. Development of this methodology is the topic of the following chapter.

4. Experiment design and pilot studies

4.0 Introduction

The aim of the experimental part of this dissertation is to investigate the claims of the Full Transfer/Full Access model of L2 acquisition (Schwartz & Sprouse 1994, 1996, outlined in Chapter 1) by comparing the performance of three groups of learners of Japanese differentiated by their L1s—English, Chinese, or Korean—with respect to their knowledge of scope interpretation in Japanese. Fulfilment of this aim necessitates the creation of a test instrument that is (i) suitable for use with L2 Japanese learners of lower proficiency as well as higher proficiency (since Full Transfer/Full Access makes claims about the initial state of L2 acquisition that may be missed if only high proficiency learners are investigated); and (ii) suitable for translation into English, Chinese, and Korean, so as to obtain control data about scope interpretation in the learners' native languages. The experimentation for the whole project is divided into three phases: Phase 1, comprising exploratory investigation of QP-QP interpretation in native Japanese; Phase 2 comprising pilot testing with English learners of Japanese as well as native English and Japanese controls; and Phrase 3, the full-scale main investigation. The present chapter details Phases 1–2; Phase 3 is the subject of Chapter 5.

This chapter is organised as follows. Section 4.1 details Phase 1: a preliminary, exploratory investigation of native Japanese judgements about a variety of QP-QP sentence types, using a picture-based acceptability judgement task. Section 4.2 details the pilot testing (Phase 2A) of a refined version of the acceptability judgement task used in Phase 1. English-speaking learners of Japanese are investigated, along with native English and Japanese control groups. Section 4.3 describes a further test of native Japanese scope interpretation (Phase 2B), using written contexts for each acceptability judgement instead of pictures. Section 4.4 summarises the findings of the development work and pilot tests, outlining how they inform the main testing carried out in Phase 3.

4.1. Phase 1: Preliminary investigation of QP-QP interpretation in native Japanese

4.1.1. Test design

The broad aim of the preliminary investigation was to find out whether the main claims of Hoji (1985) (among others) about QP-QP scope interaction in native Japanese would be exhibited by native speakers of Japanese under experimental conditions. Hoji's key claims are as follows (see Chapter 2 for detailed discussion):

152. Canonical (SOV) Japanese QP-QP sentences have only a subject-wide scope interpretation (i.e., they lack an object-wide (inverse) scope interpretation).

153. Scrambled (OSV) Japanese QP-QP sentences have both a subject-wide and object-wide interpretation.

Given that there is some disagreement about the extent to which (1) is true (i.e., about whether or not inverse scope is always absent in Japanese—see discussion of Kuno & Takami 2002, and Hayashishita 1999, 2000a, 2000b, in Chapter 2), experimental investigation of this claim was a crucial first step. In addition to this broad aim, Phase 1 was intended to investigate the effects of different quantifier types and verb types on scope interpretation in Japanese, with a view to identifying which quantifier and verb combinations produced the most robust judgements.

To fulfil these aims, ten sentence types were investigated, manipulating the following variables:[105]

154. *Variable 1: subject QP*
dareka 'someone' v. a numerically quantified noun (NumP)

155. *Variable 2: object QP*
a. *dono N-mo* 'every N' v. *subete-no N* 'all the N'
b. *daremo* 'everyone' v. *minna* 'everyone'

156. *Variable 3: verb type*
transitive v. unergative v. unaccusative

[105] As is evident from Variables 1 and 2, the subject QP was always an existential and the object, a universal. This form was used throughout all phases of experimentation, rather than a structure with a universal subject and existential object, because, as mentioned in Chapter 2, even in languages that readily allow inverse scope, such as English, the latter structure is not clearly ambiguous. For example, in *Everyone attended some seminar*, situations in which the subject-wide reading is true ('for every person, there is some seminar such that that person attended it') constitute a subset of situations in which the object-wide reading is true ('there is some seminar such that everyone attended it') if every person happened to attend the same seminar (e.g., Hornstein 1995; Liu 1997; Szabolcsi 2001). When the subject is an existential and the object a universal, there is no such entailment, so true ambiguity arises.

Variables 1 and 2 compare *wh*+QPt QPs (*dareka* 'someone', *dono N-mo* 'every N', *daremo* 'everyone') with non-*wh*+QPt QPs (NumP, *subete-no N* 'all the N', *minna* 'everyone').[106] This is in order to investigate whether the different properties of these two different types of Japanese quantifier (e.g., their ability or lack of ability to act as topics: see discussion of Tomioka (2004) in Chapter 2, Section 2.3.2.2.) affect scope interpretation. Variable 3 was included specifically to test the observation of Yatsushiro (1996) that the quantified (oblique) object of an unaccusative verb *can* take inverse scope over a subject QP in a canonical Japanese SOV sentence, unlike object QPs with transitive verbs, or locative objects with unergative verbs (see Chapter 2). The sentence types are presented in Table 9.

Table 9: Sentence types

type	subject QP	object QP	verb type	example
1	*dareka* 'someone'	*dono N-mo* 'every N'	transitive	Dareka-ga dono neko-mo nadete-iru. someone-NOM every cat-QPt is stroking 'Someone is stroking every cat.'
2	*dareka* 'someone'	*subete-no N* 'all the N'	transitive	Dareka-ga subete-no neko-o nadete-iru. someone-NOM all-GEN cat-ACC is stroking 'Someone is stroking all the cats.'
3	NumP	*dono N-mo* 'every N'	transitive	Hutari-no onnanoko-ga dono mado-mo aratta. two-GEN girl-NOM all-GEN window-QPt washed 'Two girls washed every window.'
4	NumP	*subete-no N* 'all the N'	transitive	Hutari-no onnanoko-ga subete-no mado-o aratta. two-GEN girl-NOM all-GEN window-ACC washed 'Two girls washed all the windows.'
5	*dareka* 'someone'	*daremo* 'everyone'	transitive	Dareka-ga daremo-o ai-site-iru. someone-NOM everyone-ACC loves. 'Someone loves everyone.'
6	*dareka* 'someone'	*minna* 'everyone'	transitive	Dareka-ga minna-o ai-site-iru. someone-NOM everyone-ACC loves. 'Someone loves everyone.'
7	*dareka* 'someone'	*dono N-LOC mo* 'in/on/to every N'	unerg-ative	San-wa-no tori-ga dono ki-e-mo tonda. three-CL-GEN bird-NOM every tree-LOC-QPt flew. 'Three birds flew to every tree.'
8	NumP	*dono N-LOC mo* 'in/on/to every N'	unerg-ative	Dareka-ga dono mise-ni-mo haitta. someone-NOM every shop-LOC-QPt entered. 'Someone went into every shop.'
9	*dareka* 'someone'	*dono N-LOC mo* 'in/on/to every N'	unaccus-ative	Dareka-ga dono heya-ni-mo ita. someone-NOM every room-LOC-QPt was. 'Someone was in every room.'
10	NumP	*dono N-LOC mo* 'in/on/to every N'	unaccus-ative	Ni-mai-no hanabira-ga dono isi-ni-mo otita. two-CL-GEN petal-NOM every stone-LOC-QPt fell. 'Two petals fell onto every stone.'

[106] Recall that '*wh*+QPt' quantifiers are those that comprise a *wh*-word and a quantificational particle, e.g., *daremo* 'everyone': *dare* 'who' + *mo* (particle).

The sentence types in Table 9 were investigated in their scrambled (O_iSt_iV) form as well as in the canonical (SOV) form given in the table, thus creating a fourth variable:

157. *Variable 4: word order*
SOV order v. O_iSt_iV order

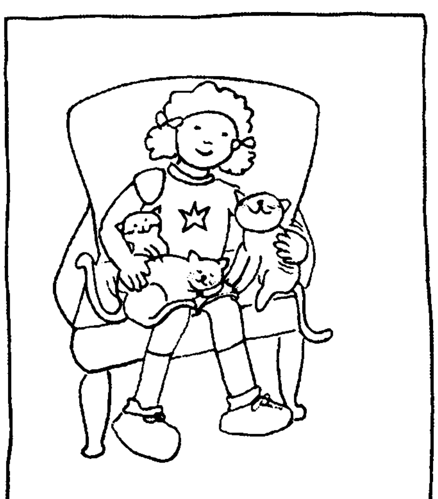
Finally, the key variable, scope interpretation, was manipulated by means of pictures provided as contexts for each test sentence.

158. *Variable 5: scope*
subject-wide scope v. object-wide scope

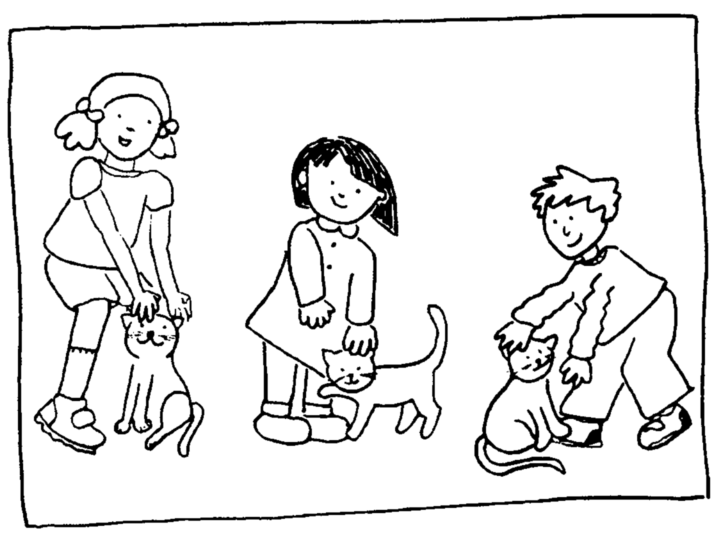
Thus each sentence could be presented in the four conditions in (159).

159. *Exemplification of the four different conditions (a, b, c, d) for Type 1 of Table 9*

<table>
<tr><th></th><th>word order</th><th>scope</th><th>sentence</th></tr>
<tr><td>a</td><td>canonical (SOV)</td><td>S>O (Picture 1, below)</td><td rowspan="2">Dareka-ga dono neko-mo nadete-iru.
someone-NOM every cat-QPt is stroking
'Someone is stroking every cat.'</td></tr>
<tr><td>b</td><td>canonical (SOV)</td><td>O>S (Picture 2, below)</td></tr>
<tr><td>c</td><td>scrambled (O_iSt_iV)</td><td>S>O (Picture 1, below)</td><td rowspan="2">Dono neko-mo dareka-ga nadete-iru.
every cat-QPt someone-NOM is stroking
'Someone is stroking every cat. (scrambled)'</td></tr>
<tr><td>d</td><td>scrambled (O_iSt_iV)</td><td>O>S (Picture 2, below)</td></tr>
</table>



Picture 1: S>O context for 159a & 159c



Picture 2: O>S context for 159b & 159d

The use of pictures to provide the scope context, rather than written (or aural) contextualisation, makes the test more suitable for lower proficiency learners of Japanese, since they will not have to read and understand contexts in their L2, in addition to tackling the actual test sentences.[107] Also, the effect of an (eventual) translation variable is minimised, since the amount of text to be translated is restricted to the test sentences. Picture contexts for judgement tasks have been used successfully in a number of previous adult L2 acquisition studies (e.g., Hirakawa 1999; Miyamoto & Yamane 1996 (see Chapter 3)).

The test battery was constructed as follows. Four sentences were created for each of Types 1–7 in Table 9, three for Type 8, and two for Types 9–10, yielding 35 basic sentences. Scrambled variants were created for three of the four basic sentences for Types 1–7, and for all of the basic sentences for Types 8–10, producing 63 test sentences. Each test sentence was then presented in both a subject-wide and object-wide scope context, yielding a total of 126 test tokens. The 126 test tokens were divided into two sets of 63 (Set 1 and Set 2). An example of how the tokens for Type 1 (subject = *dareka* 'someone'; object = *dono N-mo*, 'every N'; verb type = transitive) were divided between the two sets is provided in Table 10. The items for the nine other types were divided similarly (see Appendix 1).[108]

[107] As discussed in Chapter 3, the type of test instrument used by Dekydtspotter *et al.* (1999) (see Chapter 3), in which written contexts are presented in the learners' L1, was not considered suitable for the present project even though it would solve the problem of learners potentially failing to understand a context presented in their L2, because it creates a translation problem in studies such as this one, where the learner groups have different L1s. See also Section 4.3 for details of an exploratory pilot of a test instrument using written contexts (Phase 2B).

[108] This test construction entails a number of imbalances: there are four basic sentences for Types 1–7 but fewer for Types 8, 9 and 10; scrambled variants were presented for only three of the four basic sentences for Types 1–7; and, as Table 10 shows, the tokens for a single type were divided between two sets which were judged by two different groups of participants. The main reason for the unequal number of test tokens per type and per condition was concern that the test battery would be too long for participants to complete without losing concentration if every type had four basic sentences each presented in all four conditions. The reason for dividing the tokens of a single type between the two sets was, similarly, to reduce the monotony of the test for the participants, by ensuring that both sets contained the full variety of test types. The effects of these imbalances are discussed in Section 4.1.3.1.

Table 10: Type 1 tokens and their allocation into Sets 1 and 2[109]

item: word order; scope	sentence	set
1a: SOV; S>O	Dareka-ga dono hon-mo yonde-iru.	2
1b: SOV; O>S	someone-NOM every book-QPt is reading 'Someone is reading every book.'	1
1c: OSV; S>O	Dono hon-mo dareka-ga yonde-iru.	1
1d: OSV; O>S		2
2a: SOV; S>O	Dareka-ga kono resutoran-no dono ryouri-mo tabete-mita.	1
2b: SOV; O>S	someone-NOM this restaurant-GEN every dish-QPt eat-tried 'Someone tried every dish of this restaurant.'	2
2c: OSV; S>O	Kono resutoran-no dono ryouri-mo dareka-ga tabete-mita.	2
2d: OSV; O>S		1
3a: SOV; S>O	Dareka-ga dono neko-mo nadete-iru.	1
3b: SOV; O>S	someone-NOM every cat-QPt is stroking 'Someone is stroking every cat.'	1
4a: SOV; S>O	Dareka-ga dono bousi-mo kabutte-mita.	2
4b: SOV; O>S	someone-NOM every hat-QPt wear-tried 'Someone tried on every hat.'	2
4c: OSV; S>O	Dono bousi-mo dareka-ga kabutte-mita.	1
4d: OSV; O>S		1

Ten distractor items were added to each set (the same distractors to each set).[110] The ten distractors comprised five pairs, related in that each sentence of the pair used the same picture, or in that the sentences were the same but the pictures were different. An example is given in (160):

160. a. Hutari-no kodomo-ga Hanako-o [*e* ie-o deru toki ni] mita.
Two-GEN child-NOM Hanako-ACC house-ACC leave time at saw
'Two children$_i$ saw Hanako$_j$ as she$_j$/they$_i$ left the house.'

b. Hutari-no kodomo-ga [Hanako-ga ie-o deru toki ni] mita.
Two-GEN child-NOM Hanako-NOM house-ACC leave time at saw
'Two children saw Hanako as she left the house.'

Both (160a) and (160b) appeared with a picture of two children leaving their house and waving to Hanako who was on the other side of the street. Thus the sentence in (160a) was an appropriate description of the picture, while that in (160b) was not.

109 Glosses and translations of each 'c' and 'd' item are omitted, since these are simply the scrambled versions of the 'a' and 'b' items. The 'set' column indicates whether that test item was in Set 1 or Set 2. Thus, item 1a was in Set 2, item 1b in Set 1, etc.
110 It is acknowledged that combining only 10 distractors with 63 test items does not conform to standard recommendations that at least half the items in a judgement task be distractors (e.g., Cowan & Hatasa 1994). The fact that the actual test items represent 10 different types (Table 9), each with four conditions, mitigates this design problem, since each different type distracts from the others. Nonetheless, in later phases of experimentation the ratio of distractors to test items was increased.

Pairs of distractors were created in this way so that they would blend in with the actual test sentences, which also comprised pairs of either two different sentences with the same picture, or two identical sentences with different pictures. As well as providing distraction from the actual test items, the distractors were intended to provide a measure of whether the test participants were answering correctly. If individual participants provided unexpected judgements about the distractor items, this would be taken to indicate possible inattention, fatigue, or failure to understand how to do the test. The full list of test items and distractors is provided in Appendix 1. The order of the test items and distractors was randomised within each set.

4.1.2. Test participants

Twenty-one native speakers of Japanese participated in the investigation. All were resident in the UK at the time of testing, as undergraduate or postgraduate university students.[111] Their average length of residence in the UK and other English-speaking countries was 17 months (range: 6 months to 68 months). All spoke English as an L2 and used it daily. None were proficient in foreign languages other than English. The average age was 25 (range: 21 to 32). They were divided into two groups: Group 1 (12 participants) judged Set 1; Group 2 (9 participants) judged Set 2.

4.1.2. Procedure

Participants took the test in their respective groups. Each participant received an answer sheet listing all the test items in standard Japanese script. The picture for each test item was projected onto a screen at the front of the room using an overhead projector. Each picture was projected for 30 seconds. Participants were asked to judge how appropriately each sentence on their answer sheet matched the corresponding picture on the screen, using a three-point scale given on their answer sheets: 0 = 'not appropriate at all'; 1 = 'appropriate, but it's difficult to get that interpretation'; 2 = 'perfectly appropriate'.

The answer sheet also included pre-test and post-test questionnaires, and instructions on how to complete the test, all written in Japanese.[112] Oral instructions

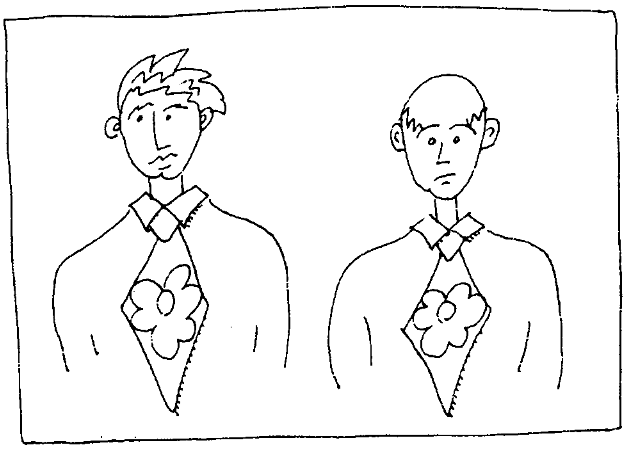
---

[111] Ideally, native intuitions about Japanese should be sought from Japanese speakers in Japan who are not subject to the daily influence of English. This ideal was fulfilled in the main studies (Phase 3).
[112] The pre-test questionnaire items asked about the participants' age, length of residence in English-speaking countries, and knowledge of other languages. The post-test items asked for comments on the length and content of the test, and the effectiveness of the instructions.

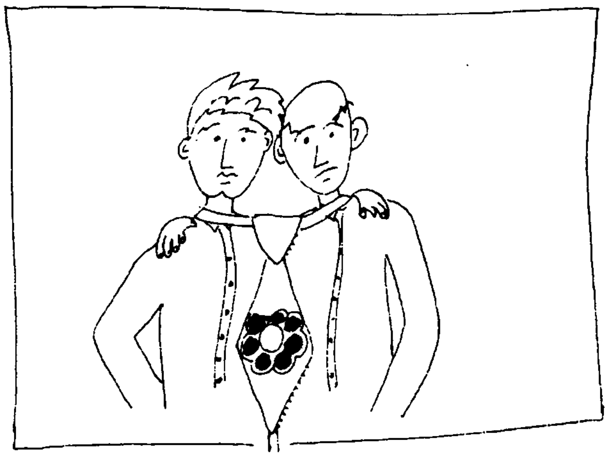
on how to complete the test were also given in Japanese, matching the instructions printed on the answer sheet. The instructions included presentation of six examples. These comprised three pairs of two examples, with the two items of each pair having the same sentence but a different picture, as illustrated in (161). (In the actual experiment, the example sentences were presented only in Japanese script; the romanisation, gloss and English translation provided here were not shown. Also, the English translation of the text in the pictures for Examples 5 and 6 was not shown.)

161.

a. *Example 1*:

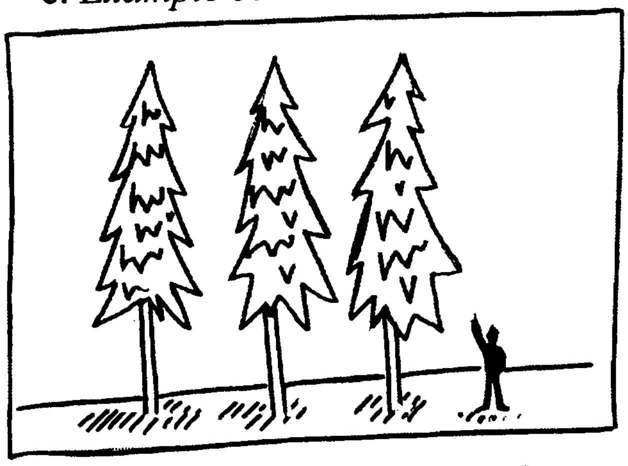


b. *Example 2*:

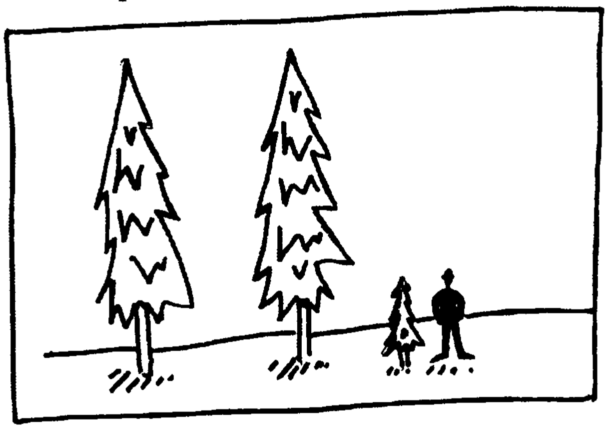


Sentence for examples 1–2:
Sumisu-san to Buraun-san-wa onazi nekutai-o site-iru.
Smith-*san* and Brown-*san*-TOP same necktie-ACC wearing.
'Mr Smith and Mr Brown are wearing the same necktie.'

c. *Example 3:*



d. *Example 4*:

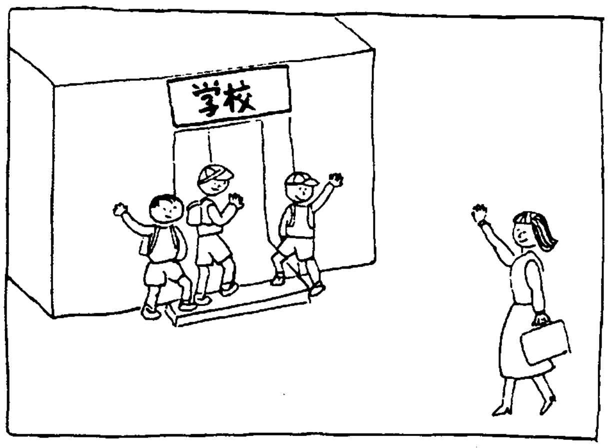


Sentence for examples 3–4:
Dono ki-mo ookii
Every tree-QPt big.
'Every tree is big.'

e. *Example 5:*

Translation (above door): School

f. *Example 6:*



Translation (above door): School

Sentence for examples 5–6:
San-nin-no otokonoko-ga sensei-o gakkou ni hairu toki ni mita.
three-CL-GEN boy-NOM teach-ACC school in enter time at saw
'Three boys saw the teacher as they/she were/was going into school.'

Just as with the actual test items, the example sentences appeared on the participants' answer sheets with the rating scale, while the example pictures were shown on the overhead projector. The instruction session included time for the participants to judge each example one by one, with discussion of the judgements and a chance to ask questions before moving on to the next example item. (In the actual test, the 73 test items were presented one after the other without discussion.)

A number of considerations informed the choice of example items. Firstly, the use of two different pictures for a single sentence was intended to prepare the participants for encountering instances of the same sentence appearing twice in the actual test, with a different picture each time. Specifically, it was hoped that the example items would prepare the participants for entertaining the possibility that a single sentence might have more than one interpretation. In the pairs comprising Examples 1–2 and 5–6 in (161), the sentence was technically appropriate for both pictures, although ratings varied considerably for Example 2, due to the unexpected nature of this picture. It was stressed that, throughout the test, there was no 'right' or 'wrong' answer—as Example 2 was intended to illustrate—thus participants should select a rating based on their intuition about the picture-sentence combination, and avoid worrying about 'right' versus 'wrong'. In the pair comprising Examples 3–4, Example 4 was intended to illustrate a clear case of the picture not being at all appropriate to the sentence. In addition, Examples 3–4 were used to draw attention to the fact that the sentence should always be considered only in the context of the picture. In other words, 'every tree' refers to every tree depicted in the picture, not every tree in the world.[113] Participants had the opportunity to ask questions about the test procedure at any stage during the instructions.

The total time taken to administer the test, including the instructions, was about 50 minutes, with the test itself taking 36.5 minutes.

4.1.2. Results

The results for the distractor items did not reveal any general- or individual-level problems with the test. The results discussed in this section thus include the judgements of all participants. The results are discussed in terms of mean ratings for each condition (exemplified in (159)) of each type (given in Table 9). Section 4.1.2.1 focuses on the effects of different subject and object QPs, keeping the verb type constant (transitive). Section 4.1.2.2. focuses on the effect of unergative verbs compared with unaccusative verbs. Due to the complexity of the test design, including the imbalances referred to in footnote 108, inferential statistics are not used for analysis of the results for Phase 1.

4.1.2.1. Results for Types 1–6

The mean appropriateness ratings (on the three point scale of 0 to 2) and standard deviations for Types 1–6 are presented in Table 11.[114]

[113] Participants' questions about precisely this point confirmed the need for this part of the explanation.

[114] Due to the way in which the test items were divided between Set 1 (12 informants) and Set 2 (9 informants), the number of judgements per token within a condition is not equal. For example, the mean rating for Type 1 on Condition (a) in Table 11 includes the ratings for two tokens judged by Group 1 (2 x 12 judgements) and two tokens judged by Group 2 (2 x 9 judgements). This factor may decrease the reliability of the results.

Table 11: Mean ratings (scale: 0–2) for S>O and O>S scope interpretations on canonical and scrambled transitive QP-QP sentences in native Japanese (standard deviations in brackets)

Type	Condition a: SOV; S>O (4 tokens/type)	Condition b: SOV; O>S (4 tokens/type)	Condition c: O_iSt_iV; S>O (3 tokens/type)	Condition d: O_iSt_iV; O>S (3 tokens/type)
1: Subj = *dareka* 'someone'; Obj = *dono N-mo* 'every N'	1.52 (0.49)	0.88 (0.57)	1.00 (0.62)	1.62 (0.45)
2: Subj = *dareka* 'someone'; Obj = *subete-no N* 'all the N'	1.76 (0.38)	0.57 (0.53)	1.50 (0.69)	1.56 (0.54)
3: Subj = NumP; Obj = *dono N-mo* 'every N'	1.97 (0.11)	0.93 (0.48)	1.98 (0.07)	1.67 (0.49)
4: Subj = NumP; Obj = *subete-no* N 'all the N'	1.93 (0.24)	0.67 (0.53)	1.95 (0.23)	1.76 (0.54)
5: Subj = *dareka* 'someone'; Obj = *daremo* 'everyone'	1.31 (0.68)	0.64 (0.60)	1.10 (0.75)	1.60 (0.63)
6: Subj = *dareka* 'someone'; Obj = *minna* 'everyone'	1.50 (0.56)	0.52 (0.42)	1.17 (0.68)	0.88 (0.72)

The overall picture presented in Table 11 indicates that an inverse (O>S) scope interpretation of canonical (SOV) Japanese QP-QP sentences (Condition (b)) is indeed considerably less acceptable than a linear (S>O) scope interpretation (Condition (a)). Mean ratings are below 0.94 for all types in Condition (b), but above 1.3 for all types in Condition (a). The smallest difference between the Condition (a) and Condition (b) means is 0.64, occurring on Type 1. On the scrambled sentences, Conditions (c) and (d), the mean ratings for both subject-wide scope (c) and object-wide scope (d) are mostly considerably higher than for object-wide scope on the canonical sentences (b). Exceptions are for Type 1 where the mean rating of 1.00 for subject-wide scope on scrambled sentences (c) is very close to the rating of 0.88 for object-wide scope on canonical sentences (b); and for Type 6 where there is less than a 0.36 point difference in the ratings for object-wide scope on canonical sentences (b) and scrambled sentences (d). Despite these two exceptions, the overall picture supports Hoji's (1985) generalisation that canonical Japanese QP-QP sentences tend to lack an object-wide interpretation, while scrambled OSV sentences allow both subject- and object-wide readings.

To consider the effect of subject QP, Types 1 and 2 (subject = *dareka* 'someone') are compared with Types 3 and 4 (subject = NumP). On the subject-wide scope Conditions, (a) and (c), the judgements with a NumP subject seem very robust, the mean rating being at least 1.93 with a standard deviation from the mean of at most only 0.24 points. When the subject is *dareka* (Types 1–2), mean ratings are somewhat lower and standard deviations higher on Conditions (a) and (c). In particular, there is a big difference on Condition (c) between Types 1 and 3. For Type 1 (subject = *dareka;* object = *dono N-mo* 'every N') the rating is just 1.00 on Condition (c), compared with 1.98 for Type 3 (subject = NumP; object = *dono N-mo* 'every N'). For the object-wide scope readings (Conditions (b) and (d)), there seems to be no great difference due to the subject QP type. The biggest difference in mean rating is just 0.2 points, occurring on Condition (d) between Type 2 and Type 4. The standard deviations range from 0.45 to 0.57 points. To summarise: judgements about subject-wide scope seem to be slightly less robust when the subject is *dareka* 'someone' compared with when the subject is a NumP, but there is no great difference between the two types of subject QP with respect to object-wide scope judgements.

The effect of different quantifiers in object position is investigated by comparing *dono N-mo* 'every N' (Types 1 and 3) with *subete no-N* 'all the N' (Types 2 and 4); and also *daremo* 'everyone' (Type 5) with *minna* 'everyone' (Type 6). Overall, no striking differences are evident. Comparing Type 1 with Type 3 and Type 2 with Type 4, the biggest differences between ratings occur on Condition (b), object-wide scope in a canonical QP-QP sentence. In both comparisons, ratings are lower when the object QP is *subete-no N* 'all the N' than when it is *dono-N-mo* 'every N' (0.57 v. 0.88 for Type 2b v. Type 1b; 0.67 v. 0.93 for Type 4b v. Type 3b). In other words, the unacceptability of object-wide scope predicted by Hoji (1985) appears to be slightly greater when the object QP is *subete-no N* 'all the N' compared with *dono N-mo* 'every N', although all four means are below the mid-point level of 1.00. However, the largest difference between means relating to the object QP occurs on Condition (c) (scrambled order with subject-wide scope), where the mean for Type 1c (object = *dono N-mo*) is just 1.00 compared with 1.50 for Type 2c (object = *subete-no N*). The low appropriateness rating on Type 1c is not expected: subject-wide scope is predicted to be available in scrambled QP-QP sentences. The degree of variation in responses to this item is also towards the high end of the range appearing in the table, with a standard deviation of 0.62 (although on Type 2c the standard deviation is actually slightly higher, at 0.69).[115]

Turning to the comparison of *daremo* 'everyone' (Type 5) with *minna* 'everyone' (Type 6), there is very little difference (maximum 0.19 points) between the two quantifiers on Conditions (a), (b) and (c). On Condition (d) (OSV order, O>S scope), however, the mean for Type 6 is depressed at only 0.88, contrasting with the high mean for Type 5, 1.60.[116]

To summarise, the lowest mean ratings for all six types occurred with an inverse scope interpretation of canonical QP-QP sentences (Condition (b)). This is as expected, given the claims (Hoji 1985, Kuroda 1970, etc.) that canonical QP-QP sentences in Japanese allow only subject-wide scope while scrambled QP-QP sentences allow both subject-wide and object-wide scope. Nonetheless, it is striking that inverse scope is not rejected outright for canonical QP-QP sentences: the participants did not choose only '0' ratings (= 'not appropriate') for tokens in Condition (b). This observation will be discussed further in Section 4.1.3. Regarding the effects of different subject QPs, judgements were more robust with numerically

[115] Since Phase 1 involved only small numbers of participants (12 or 9 informants per test token), the unexpectedly low rating for Type 1 Condition (c) could have been a chance result. Thus, space will not be devoted to in-depth exploration of why this unexpected result occurred. Briefly, however, it is interesting to note that, of the three Type 1c tokens, the ratings for 1.1c and 1.2c are particularly low:

i. Mean ratings and standard deviations for Type 1c tokens:

token	sentence	mean rating (SD)	judged by...
1.1c	Dono hon-mo dareka-ga yonde-iru. every book-QPt someone-NOM is reading 'Someone is reading every book. (*scrambled*)'	0.75 (0.75)	Group 1 (n=12)
1.2c	Kono resutoran-no dono ryouri-mo dareka-ga tabete-mita. this restaurant-GEN every dish-QPt someone-NOM eat-tried 'Someone tried every dish of this restaurant. (*scrambled*)'	0.78 (0.67)	Group 2 (n=9)
1.4c	Dono bousi-mo dareka-ga kabutte-mita. every hat-QPt someone-NOM wear-tried 'Someone tried on every hat. (*scrambled*)'	1.58 (0.51)	Group 1 (n=12)

The sentence in 1.1c and 1.2c include potentially confounding elements, as will be discussed in Section 4.1.3 with respect to the non-scrambled variants: in 1.1c, the verb is in the progressive aspect, unlike most of the other sentences of Types 1–4, where the verb is in the past tense (as in 1.4c); in 1.2c the QP object is modified by the expression *kono resutoran-no* 'of this restaurant', while no other QPs in the whole test are modified in any way. These potentially confounding factors could play a role in the unexpectedly low ratings for 1.1c and 1.2c.

[116] The low acceptability of object-wide scope for *minna* 'everyone' even in the scrambled order could indicate that *minna* is inherently collective, and does not readily allow a distributive interpretation (as asserted by Miyamoto & Yamane (1996: 501–2); but see Kuno & Takami (2002, Chapter 4) for examples of *minna* used distributively). Further investigation of the behaviour of *minna* compared with other Japanese words meaning 'everyone' would shed light on this hypothesis. Due to the need to narrow the scope of the present project (see Section 4.1.3), the question is not pursued further here.

quantified nouns as the subject than with *dareka* 'someone'. No striking effects of object QP were evident, although on Condition (b) (SOV order, object-wide scope), ratings for *subete-no N* (Types 2 and 4) were slightly lower than for *dono N-mo* (Types 1 and 3), suggesting that inverse scope might be slightly less readily available when the object QP is *subete-no N*.

4.1.2.2. Results for Types 7–10

Types 7–10 investigate scope interpretation with unergative and unaccusative verbs. The mean appropriateness ratings are presented in Table 12.[117]

Table 12: Mean ratings (scale: 0–2) for S>O and O>S scope interpretations on canonical and scrambled unergative and unaccusative QP-QP sentences in native Japanese (standard deviations in brackets)

Type (Obj = *dono N-LOC-mo* 'in/on/to every N' throughout)	**Condition a:** **SOV; S>O**	**Condition b:** **SOV; O>S**	**Condition c:** **O_iSt_iV; S>O**	**Condition d:** **O_iSt_iV; O>S**
7: subj = *dareka* 'someone'; verb = unergative	1.62 (0.59) *(4 tokens)*	0.88 (0.59) *(4 tokens)*	0.88 (0.72) *(3 tokens)*	1.95 (0.15) *(3 tokens)*
8: subj = NumP; verb = unergative	1.88 (0.27) *(3 tokens)*	0.95 (0.69) *(3 tokens)*	1.26 (0.61) *(3 tokens)*	1.88 (0.27) *(3 tokens)*
9: subj = *dareka* 'someone'; verb = unaccusative	0.76 (0.83) *(2 tokens)*	1.76 (0.44) *(2 tokens)*	0.38 (0.59) *(2 tokens)*	1.95 (0.22) *(2 tokens)*
10: subj = NumP; verb = unaccusative	1.19 (0.75) *(2 tokens)*	1.52 (0.68) *(2 tokens)*	0.91 (0.63) *(2 tokens)*	1.95 (0.22) *(2 tokens)*

The ratings for unergative verbs (Types 7 and 8) are similar to those for transitive verbs (Table 11): in canonical QP-QP sentences, subject-wide scope is readily available (Condition (a)), while the object-wide scope ratings are below 1.00 (Condition (b)). In the scrambled variants, object-wide scope is readily available (Condition (d)), with ratings over 1.85. The ratings for subject-wide scope (Condition (c)), are lower: with *dareka* 'someone' as the subject (Type 7), the mean rating is only 0.88, while with a numerically quantified subject (Type 8) it is slightly higher at 1.26.

The ratings for the unaccusative verb types (Types 9 and 10) in the non-scrambled conditions (Conditions (a) and (b) are strikingly different: subject-wide scope receives rather low mean ratings (0.76 for Type 9, Condition (a), 1.19 for Type

[117] As with Types 1–6, the tokens for each type in Table 12 were divided between Set 1 (12 informants) and Set 2 (9 informants). Thus, the mean ratings for Type 7 on Condition (a), for example, include two tokens judged by 12 informants and two tokens judged by a separate group of 9 informants.

10, Condition (a)). However, object-wide scope receives mean ratings of greater than 1.5—close to the 'perfectly appropriate' end of the judgement scale. Thus, Yatsushiro's (1996) observation that inverse scope is available with unaccusative verbs in SOV sentences is borne out. However, the status of subject-wide (S>O) scope with these verbs is not so clear. The standard deviations for Condition (a) are large (≥0.75), indicating a considerable degree of inter-participant variation. Considering there were only two tokens for each condition for Types 9 and 10, the most conservative interpretation of the Condition (a) results is that they simply show that judgements are varied. The data are not clear enough to support or refute a claim that subject-wide scope is lacking in SOV QP-QP sentences with unaccusative verbs.

Moving on to the scrambled sentences with unaccusative verbs, subject-wide scope (Condition (c)) was essentially unavailable when the subject was *dareka* 'someone' (rating = 0.38) and was less than marginally available when the subject was a numerically quantified noun (rating = 0.91). By contrast, object-wide scope (Condition (d)) was highly acceptable, with ratings of 1.95 for both Types 9 and 10.

To summarise, there is a clear difference between unergative and unaccusative verbs with regard to the availability of object-wide scope in SO-LOCV QP-QP sentences: unaccusative verbs readily allow inverse scope, while unergative verbs do not. Regarding the effect of varying the subject QP, the availability of subject-wide scope appears to decrease with unaccusative verbs (Types 9 and 10) when the subject is *dareka* 'someone' compared with a NumP. With unergative verbs, this pattern of acceptability (subject-wide scope being less readily available with *dareka* than with a NumP) obtains in just the scrambled condition, (c).

4.1.3. Discussion: Implications of the Phase 1 findings for the Phase 2 design

The results discussed in the two preceding sections broadly confirm the claims of the theoretical literature (Hoji 1985; Kuroda 1970; Yatsushiro 1996 for unaccusatives): canonical SOV Japanese QP-QP sentences were not readily acceptable in an inverse scope context, except with an unaccusative verb. Scrambled OSV QP-QP sentences, however, were acceptable in both subject-wide and object-wide contexts (with some exceptions, particularly with unaccusative verbs). The fact that the predicted effects of varying the scope context (S>O v. O>S) and the subject and object order (canonical SOV v. scrambled OSV) were successfully detected by the test instrument indicates that, despite the imbalances in the materials design, the test instrument did indeed measure what it was designed to measure. Thus it was decided to continue working with the same basic design of QP-QP sentences judged in the context of pictures. Needless to say, however, the findings of this initial investigation informed numerous changes to all aspects of the design eventually used in the pilot study and the main study. In particular, attention was paid to trying to maximise the reliability of the results. The crucial detail of Japanese QP-QP scope interpretation in the context of the present dissertation is the claim that Japanese lacks inverse scope. The findings of this exploratory study showed that inverse scope was considerably less readily available than subject-wide scope on SOV transitive QP-QP sentences. However, as already noted, inverse scope was not absolutely ruled out: mean ratings for inverse scope on SOV transitive QP-QP sentences ranged from 0.52 to 0.93 on the scale of 0 to 2, yet the theoretical claims of Hoji (1985) would predict ratings approaching '0'. It is not clear from this initial study whether these ratings represent a reality whereby the inverse scope interpretation is in fact available in Japanese to a limited extent, or whether aspects of the experiment design served to make the ratings for inverse scope higher than they should be. This section will therefore detail how the test battery and the procedure were revised, with the aim of maximising the reliability of subsequent experiments.

4.1.3.1. Test battery

There were a number of imbalances in the test construction (as already mentioned in footnote 108): there were four basic sentences for Types 1–7 but fewer for Types 8, 9 and 10; scrambled variants were presented for only three of the four basic sentences for Types 1–7; and the tokens for a single type were divided between two sets which were judged by two different groups of participants. This test design emerged as a result of trying to balance the need for a test that was not too long or monotonous with the desire to investigate as many sentence types as possible. However, the resulting imbalances affect the analysis of the results. For example, responses to four tokens of Type 7 had to be compared with responses to only three tokens of Type 8 when investigating the effect of subject and object QP in canonical sentences with unergative verbs. This decreases the reliability of the results. Thus, an obvious step towards increasing reliability for the subsequent phases of experimentation involves avoiding the test battery imbalances of Phase 1. Specifically, the test battery should include an equal number of tokens per type, and sentences of the same type should not be judged by two different groups of participants.

It was decided that the Phase 2 test set should be short enough to be judged by a single set of participants. The Phase 1 post-test questionnaire responses about the test duration were consulted, in order to inform the decision about how long the Phase 2 test could reasonably take. Eleven of the 21 Phase 1 participants indicated that the test duration (36.5 minutes, excluding explanation time) had been too long, the remaining 10 responding that it was fine. This was interpreted to indicate that the test duration for Phase 2 should be reduced, especially considering that Phase 2 and Phase 3 participants would include learners of Japanese, for whom the burden of taking the test in their L2 can be assumed to be greater than for participants taking it in their native language. It was decided that the test should take no longer than 30 minutes (excluding explanation time), and that Phase 2 participants should also be asked to comment on the length of the test.

A goal for Phase 2 was to focus on those Phase 1 sentence types that had produced the most robust native Japanese judgements, in order to see if those judgements would be replicated by a different native Japanese group (thereby providing a further measure of the reliability of the test instrument). Standard deviations (Table 11 and Table 12) were thus examined in order to find out which of the types in Phase 1 were judged with the lowest levels of variation. These were found to be Type 3 (subject = NumP, object = *dono N-mo* 'every N', verb = transitive) and Type 4 (as Type 3, but with *subete-no N* 'all the N' as object). It was decided to continue working with these types, and also with Types 1 and 2, which differ minimally from Types 3 and 4, respectively, in that the subject is *dareka* 'someone' rather than a numerically quantified NP.[118] Consequently, the 'verb type' variable was removed. In addition, it was decided to focus on the SOV sentence order, and drop the scrambled sentences from the Phase 2 test battery.

In order to inform the creation of test items for Phase 2, the mean ratings and standard deviations of the individual test tokens of Types 1–4 were examined, with the aim of identifying any tokens that produced less robust judgements than others. This examination confirmed that for each individual test sentence used in Types 1–4, the mean ratings for Condition (a) (subject-wide scope) were higher than the mean

[118] The standard deviations for Types 1 and 2 on at least the non-scrambled conditions (Conditions (a) and (b)) were also generally slightly lower than for Types 5–10.

ratings for Condition (b) (object-wide scope), as predicted by the theoretical literature, and as expected from the overall mean ratings by type given in Table 11.[119] It was decided, therefore, to continue working with these test items, but to look for ways in which the sentences themselves or the pictures might be improved.

Considering, firstly, the structure of the sentences, the use of the progressive aspect in two sentences of Type 1 and Type 2 compared with the past tense for all other sentences of Types 1–4 could have affected the degree of variation in ratings for those sentences. Table 13 shows that in Condition (b) (the key test condition: SOV order with object-wide scope) on Type 2, the two sentences in which the verb was in the progressive form (2.1b and 2.3b in Table 13) yielded higher standard deviations (i.e., more variation in ratings) than the two sentences in which the verb was in the past form (2.2b and 2.4b). (The indices '2.1b', etc., are the indices of these test items in the full test battery given in Appendix 1. '2.1b' indicates 'Type 2, sentence 1, Condition (b)'.)

Table 13: Mean ratings and standard deviations for Type 2 sentences in Condition (b) (SOV order, object-wide scope)

index	sentence	mean rating	SD
2.1b	Dareka-ga subete-no hon-o yonde-iru. someone-NOM all-GEN book-ACC is reading 'Someone is reading all the books.'	0.44	0.73
2.2b	Dareka-ga kono resutoran-no subete-no ryouri-o tabete-mita. someone-NOM this restaurant-GEN all-GEN dish-ACC eat-tried 'Someone tried all the dishes of this restaurant.'	0.33	0.49
2.3b	Dareka-ga subete-no neko-o nadete-iru. someone-NOM all-GEN cat-ACC is stroking 'Someone is stroking all the cats.'	1.11	0.78
2.4b	Dareka-ga subete-no bousi-o kabutte-mita. someone-NOM all-GEN hat-ACC wear-tried 'Someone tried on all the hats.'	0.5	0.52

[119] The mean rating on each S>O scope item (Condition (a)) was at least 0.5 point higher (on the scale of 0–2) than that of the corresponding O>S scope items (Condition (b)), with one exception: the Type 1 sentence in (i) yielded a mean rating of 1.33 in the S>O condition (a) and 1.00 in the O>S condition (b) (difference = 0.33 points).

i. Dareka-ga dono hon-mo yonde-iru.
someone-NOM every book-QPt is reading
'Someone is reading all the books.'

As will be shown, problems were identified with both the picture and the sentence structure of this item. These may have been the reason for the lack of clear differentiation between subject-wide and object-wide scope acceptability.

It cannot be claimed that the higher standard deviations for 2.1b and 2.3b (0.73 and 0.78, respectively) arose *because of* the progressive aspect versus past tense difference. The higher variation in 2.1b and 2.3b could also be due to differences between the two test groups, as 2.1b and 2.3b were both judged by Group 2, while 2.2b and 2.4b were judged by Group 1. Alternatively, the result could be due to chance, because the numbers of informants per group (12 and 9) were rather small. Nonetheless, the pattern of lower variation on past tense items, higher variation on progressive aspect items draws attention to the potential for the tense of the verb to be a confounding factor in analysis of the results.[120] Thus, it was decided to present all sentences in the past tense in Phase 2.

One further change to ensure that all Phase 2 sentences had the same basic structure was the removal of the modifying expression *kono resutoran-no* 'of this restaurant' in the sentence in (162), used in Type 1 (object = *dono N-mo* 'every N') and Type 2 (object = *subete-no N* 'all the N'):

162.	Dareka-ga	kono resutoran-no	(i) dono ryouri-mo (ii) subete-no ryouri-o	tabete-mita.
	someone-NOM	this restaurant-GEN	(i) every dish-QPt (ii) all-GEN dish-ACC	eat-tried

'Someone tried every dish/all the dishes of this restaurant.'

No other test sentence for Types 1–4 contained modified QPs. The inclusion of the pre-QP modification could again be a confounding factor affecting judgements about the sentence.[121]

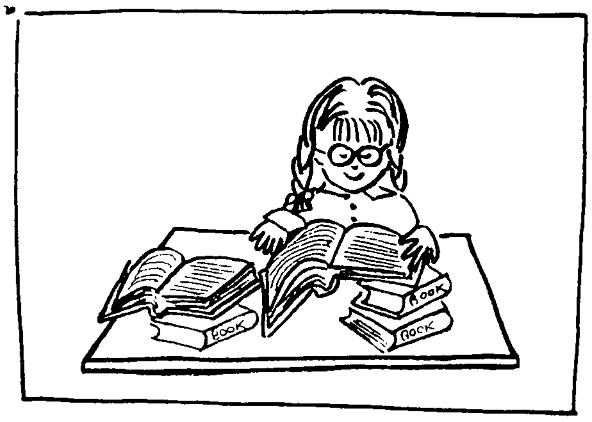
The reason that some of the Phase 1 test sentences were constructed with the progressive aspect was that simple pictures of the activities in question (*stroke (a cat)* and *read (a book)*) lent themselves more readily to a progressive interpretation than a past tense interpretation. This is exemplified in (163): (163b) was used for item 2.1b in Table 13 and (163a) for the subject-wide variant of the same sentence (2.1a).

[120] On Type 1, the standard deviations for Condition (b) on the progressive aspect items (0.75 on 1.1b, 0.74 on 1.3b) did not differ greatly from those on the past tense items (0.67 on 1.2b, 0.71 on 1.4b), although here, too, the lowest standard deviations were indeed on the past tense items (1.2b, 1.4b) and the highest on the progressive aspect items.

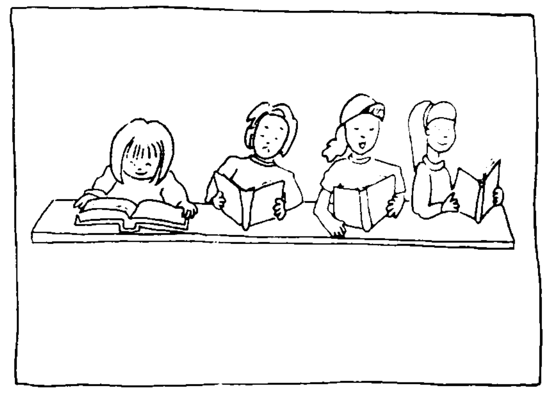
[121] In fact, the mean ratings for this sentence in Types 1 and 2, Conditions (a) and (b), do not suggest that the sentence was problematic: in each case the subject-wide scope interpretation was rated highly (1.83 for token 1.2a, 1.78 for token 2.2a), while the object-wide scope interpretation received low ratings (0.78 for 1.2b, 0.33 for 2.2b).

163.

a.



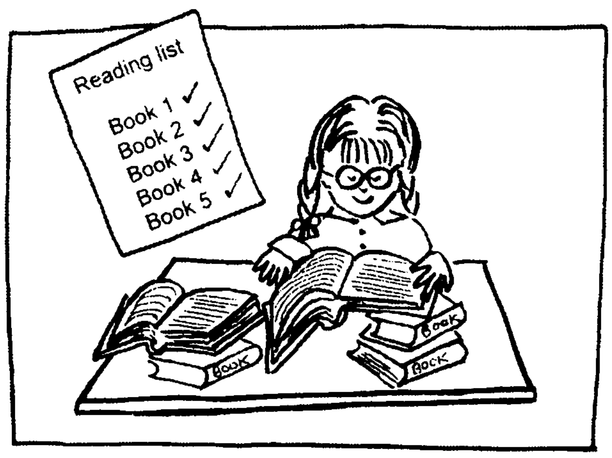
b.



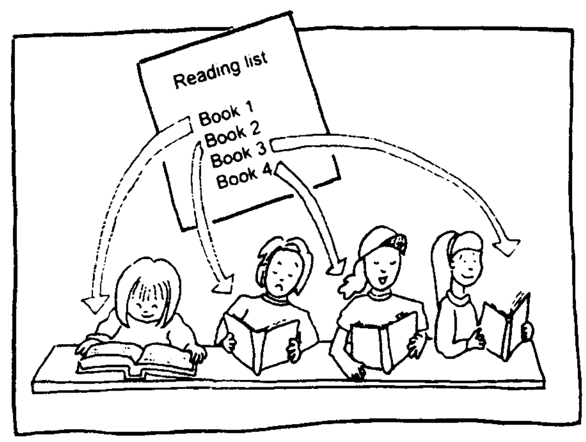
The most obvious interpretation of the pictures in (163) is that people are (in the process of) reading books. If participants are asked to judge a past-tense sentence, *Someone read every book/all the books*, in the context of these pictures, they might rule it out on the grounds that the tense is wrong. To try to solve this problem, it was decided to state in the instructions for taking the test that all pictures in the test depict events that took place 'yesterday', and to emphasise this in the example items—which would also all be changed to the past tense—prior to taking the test. In addition, pictures were amended wherever possible, with the aim of making a past tense reading more plausible. The pictures in (163) were thus revised for Phase 2, to include a 'Reading List' indicating a specific set of books, as shown in (164).

164.

a.



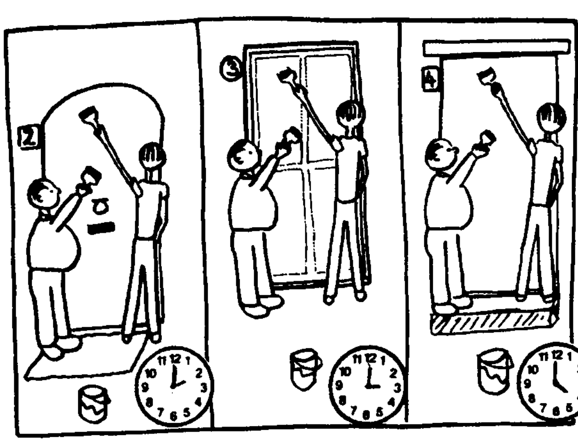
b.



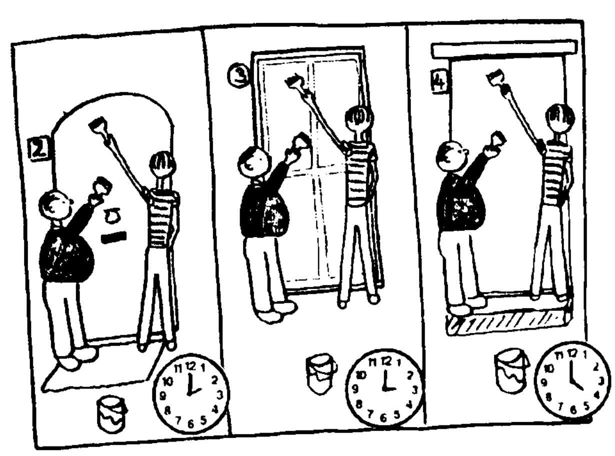
In (164a), the tick beside each book on the list is intended to indicate that the girl has read (at least some of) each one. Similarly in (164b), the association of a particular book with each person is intended to contribute to the idea of each person having read (at least some of) their designated book. In conjunction with the instruction that the pictures all depict events that happened the day before, it was hoped that this modification to the pictures would increase the plausibility of a past tense reading. In the version of the test to be used with native speakers of Japanese, the English words 'Reading List' were replaced by the Japanese equivalent, *sankou-bunken*, and the ticks were replaced with circles, which are more usual in a Japanese context. (Since all participants were university students, it was assumed that the concept of a reading list would be familiar.)

Each picture was thus examined in order to identify (i) further improvements that might increase the plausibility of the past tense reading, or avoid unintended interpretations of the picture, and (ii) culture-specific images which might require 'versioning' for participants from different cultures (such as the ticks in (164a), mentioned above). A major change to the pictures was the inclusion of more colour in order to help clarify when the same individual was depicted more than once, or to distinguish between different individuals who might potentially be assumed to be the same. The need for further clarification of some of the complicated pictures was identified by a post-test questionnaire comment that suggested using more colour, and by talking to the participants after they had completed the test. An example of a picture to which colour was added is that in (165a), which was used for the subject-wide reading of *Hutari-no otoko-ga dono doa-mo nutta* 'Two men painted every door' in both Phase 1 and Phase 2 (and, eventually, Phase 3). In Phase 1, the picture was not coloured (165a): participants were expected to judge from the body shapes that the same two men are painting each door. In the version used for Phase 2 (165b), this judgement was further facilitated by the man on the left of each door having a blue sweater while the man on the right had a red and white striped T-shirt.

165. a. *Phase 1 (no colour)*:



b. *Phase 2 & 3 (colour on shirts)*:



A number of pictures were touched up in this way for Phase 2. In addition, the pictures for *Hutari-no kangohu-ga dono kanzya-mo kanbyou-sita* 'Two nurses looked after every patient' were completely re-drawn (compare the pictures for 3.4a and 3.4b in Appendix 1B with 12a and 12b in Appendix 2A). The design of the originals used in Phase 1 was too complex, so it was not clear enough who was a patient and who was a nurse. Also, the picture for (3.4b) contained an error in that one of the nurses was male. The Japanese word used for nurse, 看護婦, *kangohu*, contains a character (婦 *hu*) that indicates that the noun is feminine: thus all nurses in the picture should have been female.[122]

The distractor sentences and their corresponding pictures were also examined and revised for Phase 2. Six of the ten distractors in Phase 1 were biclausal sentences, in contrast to the monoclausal sentences of the other 130 items in the test battery. For Phase 2, the distractors were revised to be monoclausal SOV sentences in the past tense, like the test items, so that the participants' attention is not potentially drawn to the distractors due to their different structure. In addition, the Phase 2 distractors had either a subject or object QP. Finally, the ratio of distractors to test items was increased, as this ratio was very low, 1:6.3, in Phase 1. Details of the revised distractor set are presented in Section 4.2.

This section has detailed a number of changes made to the Phase 1 test battery in order to adapt it for Phase 2 and maximise the reliability of the experimentation. First, the problem of potentially compromised reliability due to imbalances in the Phase 1 test battery was addressed: for Phase 2 the test battery was created with equal numbers of tokens per type, and it was short enough to be completed in less than 30 minutes without division into two sets. Second, the sentence types for investigation in Phase 2 were identified as Types 1, 2, 3, and 4 in their non-scrambled form. These types were selected on the basis of the relatively low standard deviations from their mean ratings, which indicate that judgements about these sentences were generally robust. Third, each individual token of Types 1–4 was examined with respect to its mean rating and standard deviation, in order to discover whether any individual token produced anomalous ratings. It was discovered that some sentences had verbs in the progressive aspect while the

---

[122] I realised from talking to participants after the Phase 1 and Phase 2 tests that, even if the Japanese word *kangohu* '(female) nurse' were gender-neutral, all images in the pictures should be as stereotypical as possible, to avoid arousing suspicions that a test sentence is 'not appropriate' because, for example, a child in the picture looks boyish, but the sentence says 'girls'. Thus (contra current guidelines for academic and educational writing), girls should be portrayed wearing skirts and ribbons, boys wearing caps and shorts, nurses as females, police officers as males, and so on.

majority had verbs in the past tense, and that the progressive tokens yielded greater inter-participant variation than the past tense tokens. Consequently, Phase 2 test sentences were created with the uniform structure of [QP-NOM QP-ACC V.PAST]. Fourthly, the pictures for the Type 1–4 test items were re-examined to identify how their clarity could be improved in order to minimise any judgements of a picture-sentence mismatch on the basis of a misunderstanding of the picture. As a result, colour was added to most pictures for Phase 2, and a number of pictures were modified or redrawn. Finally, it was noted that distractor items should have the same [SOV.PAST] structure as the test items, and that the ratio of distractor items to test items should be increased.

4.1.3.2. Test procedure

In addition to revising the test battery, as described in the previous section, a number of changes were made to the test procedure for Phase 2. Firstly, the rating scale was changed from three points (0 to 2) to five points (–2 to +2). This was motivated by evidence that scales with more than three points are statistically more reliable (Sorace 1996: 398). The second important change was to remove the test sentences from the answer sheet and present them on the screen with the picture. Details of this change comprise the topic of the rest of this section.

The motivation to remove the test sentences from the answer sheet came from observing the Phase 1 participants' behaviour while taking the test. It was clear that, after the first few test items, participants started to notice the similarities between the test items and to look back at their previous ratings for similar sentences. In so doing, they may have been tempted to change answers retrospectively, which seriously threatens the reliability of the ratings, since any retrospective changes would have been made without even being able to see the relevant picture. Removing the sentences from the answer sheet (leaving only the judgement scale) was thus intended to reduce the chance of any retrospective changes. It also removes the possibility of participants rapidly reading through all the test sentences and forming preconceptions about their interpretations before seeing the pictures. Thus the chance of judgements reflecting participants' on-line intuitions rather than premeditated judgements are increased.

It was decided that, for each test item in Phase 2, the picture would be projected onto the screen without the test sentence for 10 seconds, then the test sentence would be revealed beneath the picture for the remaining 20 seconds of the test item. Judgements about each item would therefore have to be made within the 20 seconds of viewing each picture and sentence together. In order to facilitate reading of the test sentences, particularly for the learners, it was decided that every test sentence should also be presented aurally at the same time as it was revealed in written form on the screen.

To summarise, the main changes to the test procedure identified on the basis of Phase 1, were to change the rating scale from a three-point to a five-point scale, to take the test sentences off the answer sheet and present them only on the overhead projection screen, and to include aural presentation of the test sentences. These changes were intended to increase the reliability of the data collected in Phase 2.

4.2. Phase 2A: A pilot study of QP-QP interpretation in L2 Japanese[123]

The pilot study had three main goals: (i) to find out whether the results of the preliminary test (Phase 1) with native speakers of Japanese would be replicated, thereby confirming the reliability of the test; (ii) to find out whether native English speakers would indeed accept inverse scope on an English version of the test, thereby further confirming the sensitivity of the test instrument; (iii) to trial the test instrument with English-speaking learners of Japanese and test two predictions (166a–b) based on Full Transfer/Full Access:

166. a. *Prediction 1: (L1 Transfer)*
Due to L1 transfer, learners whose L1 is English will allow non-target-like ambiguity in Japanese doubly-quantified sentences.

b. *Prediction 2: (L2 Access to UG)*
Due to L2 access to UG, English-speaking learners of Japanese *can* acquire Japanese scope rigidity (despite underdetermination by the evidence).

4.2.1. Test design

The test instrument was a modified version of the acceptability judgement task used in Phase 1, the modifications being based on the Phase 1 findings, as outlined in Section 4.1.3. Three variables were manipulated:

167. *Variable 1: subject QP*
dareka 'someone' v. a numerically quantified noun (NumP)

[123] The results of this pilot study have already been reported in Marsden (2003).

168. *Variable 2: object QP*
dono N-mo 'every N' v. *subete-no N* 'all the N'

169. *Variable 3: scope*
subject-wide scope v. object-wide scope

Eight test item types were created with the structure [subject-QP object-QP verb.PAST], as shown in Table 14.

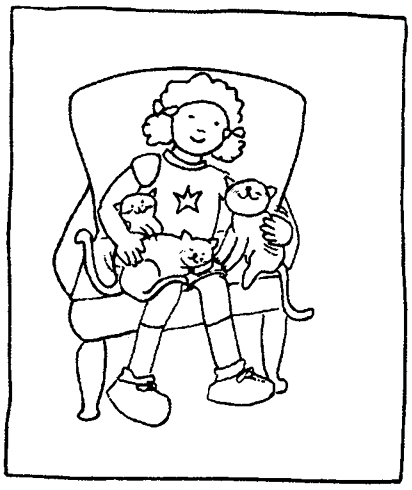
Table 14: Test item types

type	subject QP	object QP	scope	example
1a	*dareka* 'someone'	*dono-N mo* 'every N'	S>O	Dareka-ga dono neko-mo nadeta. someone-NOM every cat-QPt stroked 'Someone stroked every cat.'
1b			O>S	
2a		*subete-no N* 'all the N'	S>O	Dareka-ga subete-no hon-o yonda. someone-NOM all-GEN book-ACC read 'Someone read all the books.'
2b			O>S	
3a	NumP	*dono-N mo* 'every N'	S>O	Sannin-no onnanoko-ga dono tako-mo ageta. three-GEN girl-NOM every kite-QPt flew 'Three girls flew every kite.'
3b			O>S	
4a		*subete-no N* 'all the N'	S>O	Hutari-no onnanoko-ga subete-no mado-o aratta. two-GEN girl-NOM all-GEN window-ACC washed 'Two girls washed all the windows.'
4b			O>S	

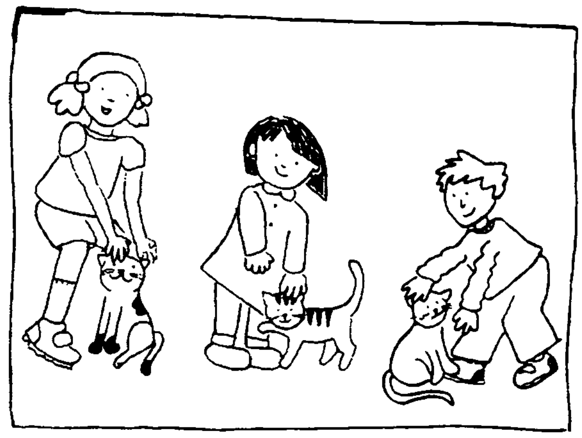
An English version of the test was created for the native English control group. For examples of the English sentences, see the translations in Table 14.[124] As in Phase 1, the scope variable was manipulated by pictures depicting either subject-wide or object-wide interpretations of each sentence. To illustrate, the pictures for 1a and 1b in Table 14 are given in (170a–b). (The picture in (170a) is identical to the one used for (159a) and (159c) in Phase 1; the picture in (170b) differs from that used in (159b) and (159d) in that markings were added to the cats in order to clarify that each child is stroking a different cat.)

[124] Throughout the remainder of this dissertation, the English test sentences can always be assumed to be the same as the translations given for the Japanese test sentences, unless otherwise indicated.

170. a. *S>O scope context:*



b. *O>S scope context:*



Four tokens of each type were created, yielding 32 test items. Fourteen distractor items were added to these, making a test battery of 46 items. Care was taken to select vocabulary that was likely to be familiar to the learners of Japanese.

The distractor items also had the structure [subject object verb.PAST], although some additionally contained a floating quantifier, as illustrated in (171).

171. Kodomo-ga hutari-tomo Suuuzan-o mita.
child-nom both Susan-acc saw
'Both children saw Susan.'

To further blend in with the test items, some of the distractors had a subject or object QP. Nine of the 14 distractors were intended to not match the picture they were presented with; the remaining five each matched their respective pictures. Consideration of the English version of the test was behind the decision to include more 'unacceptable' than 'acceptable' picture-sentence pairs among the distractors: in the English version, the object-wide (O>S) scope types with *every* as the object quantifier (1b and 3b in Table 14) were expected to be judged acceptable, unlike in Japanese where all the object-wide scope types (1b, 2b, 3b and 4b) are unacceptable. Thus, the greater number of 'unacceptable' distractors was intended to counter, at least slightly, the larger number of 'acceptable' test items in English. All distractor items were intended to be unambiguous.

The first two and last two items on the test were distractors. Distractors were used for the first two items in order to give the participants a short 'warm-up' in the test, and for the last two items as an attention check (because, on approaching the end of the test, participants may stop paying attention and answer randomly as they stop concentrating). The remaining ten distractors were mixed randomly with the test items to create the test battery. The distractor results were used to identify any participants whose data should be excluded from the analysis due to possible misunderstanding of the task (see Section 4.2.4.1).

Two presentation orders were created for the test, Order 2 being the reverse of Order 1. The full test battery is given in Appendix 2A.

4.2.2. Participants

Eighteen English-speaking learners of Japanese ('EJ'), 33 native Japanese speakers ('JJ') and 29 native English speakers ('EE') participated in the experiment. Details of the participants are given in Table 15.

Table 15: Participants

group	no.	age		details
		range	mean	
EJ	18	21–34	24	All but one were final-year students of Japanese at UK universities, having had approximately 700–800 hours of instruction and having spent at least nine months (average 1 year 4 months) in Japan. The one exception had had approximately 180 hours of university-level Japanese instruction followed by three years living in Japan.
JJ	33	18–29	19	Mainly 1st-year undergraduates recently arrived at a UK-based Japanese university.
EE	29	18–51	21	Mainly British undergraduate and postgraduate students.

The learners' proficiency level was judged to be 'high intermediate/advanced' on the grounds that this is the level usually attained by learners in their final year of Japanese studies at UK universities. However, no independent measure of proficiency was made for the pilot test. (See Chapter 5 for details of the proficiency test used in the main study.) The learners of Japanese had all had experience of studying foreign languages other than Japanese prior to beginning their study of Japanese. All had begun learning Japanese as adults. Ideally, learners for whom Japanese was the first L2 would have been selected. However, in the UK, such learners are very rare. Therefore it was not possible to control for this variable in the pilot study or the main study.

4.2.3. Procedure

Section 4.1.3.2. explained how the procedure for the pilot test was developed from the procedure for Phase 1. The present section details the procedure without reiteration of how it was developed.

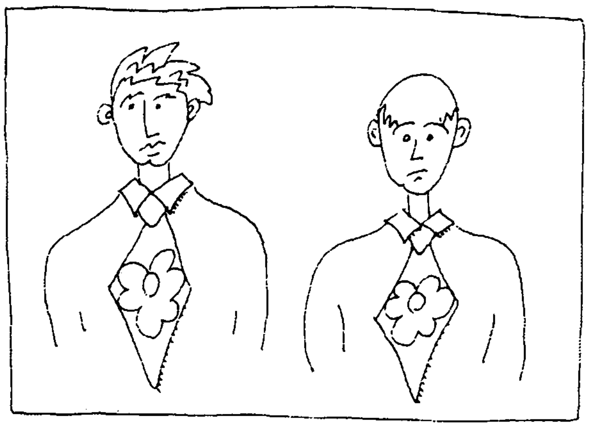
Participants were asked to judge whether each test sentence matched the picture shown for that sentence, and to indicate their answers on a five-point scale of –2 to +2, where '–2' meant 'No, not at all' (i.e., the sentence does not match the picture at all), and '+2' meant 'Yes, [it matches] perfectly'. The sentences and pictures were projected onto a screen at the front of the room: they did not appear on the participants' answer sheets. Each picture was projected without the sentence for 10 seconds, then the sentence was revealed below the picture for a further 20 seconds. At the same time as the sentence was revealed, it was also presented aurally.[125] The aural presentation was included to facilitate the learners' understanding, should they have encountered any difficulty in reading the sentences. In addition, in the Japanese version of the test, *furigana* (phonetic characters) were used above all but the most basic of the *kanji* (ideographs), so as to further facilitate ease of reading and comprehension. The test took 23 minutes to run.

As in Phase 1, the answer sheet included pre-test and post-test questionnaires, and instructions on how to complete the test. A sample answer sheet is presented in Appendix 2B. Prior to running the test itself, oral instructions were given on how to complete the test. These were based on the instructions printed on the answer sheet. The instructions (both oral and written) were given in English to the L2 Japanese and the native English groups, and in Japanese to the native Japanese group. Six example test items were presented in the instructions, using the sentences and pictures in (172). (In the actual experiment, the example sentences were presented only in Japanese script in the Japanese version and English in the English version; the romanisation, gloss and English translation provided here were not shown.)

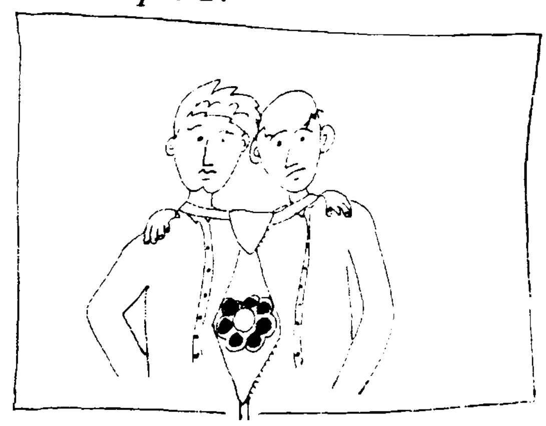
[125] In the English version of the test, the researcher read out the sentences. For the Japanese version, an audiotape recorded by a native speaker of Japanese was usually used, although for one group of native Japanese participants, the sentences were read out by a native speaker of Japanese who was assisting the researcher for that test session.

172.

a. *Example 1*:

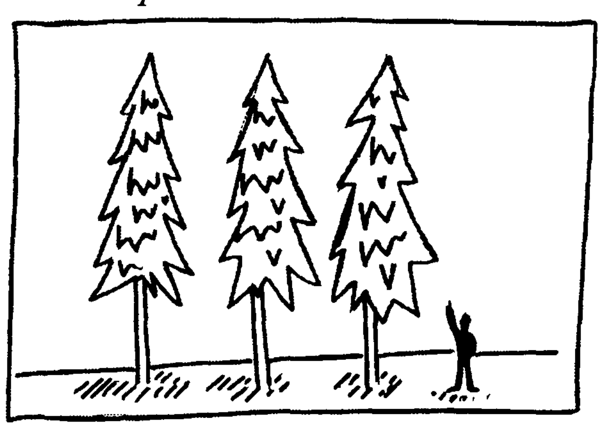


b. *Example 2*:

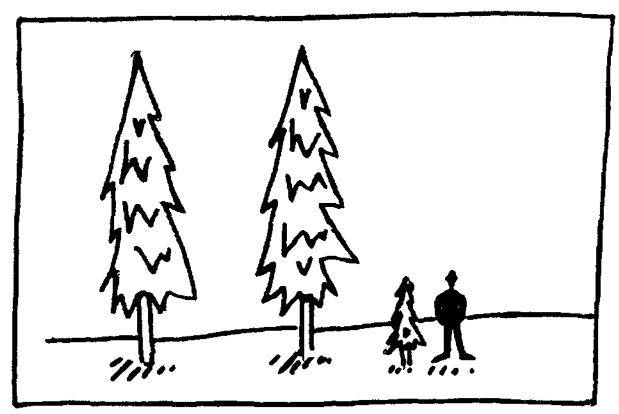


Sentence for examples 1–2:
Hutari-no otoko-ga onazi nekutai-o sita.
two-GEN man-NOM same necktie-ACC wore.
'The two men wore the same necktie.'

c. *Example 3*:

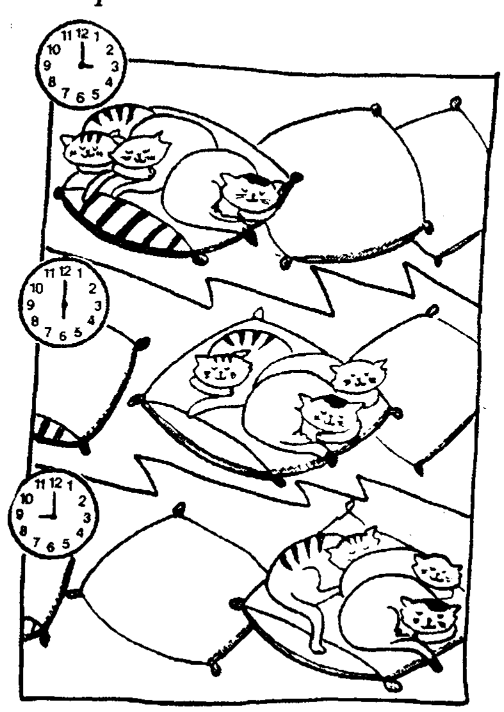


d. *Example 4*:

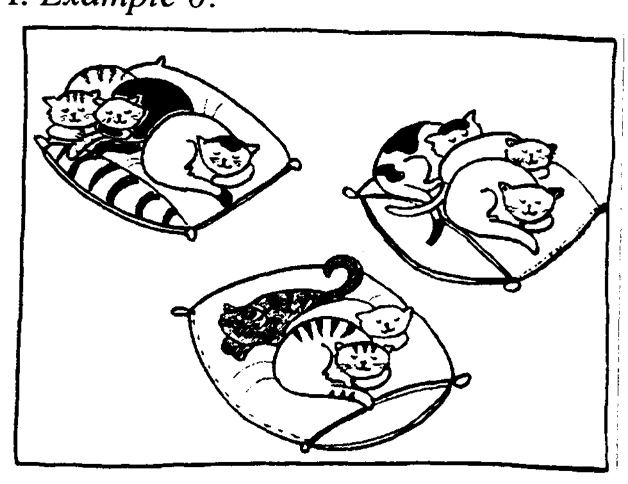


Sentence for examples 3–4:
Dono ki-mo ooki-katta
Every tree-QPt big-PAST.
'Every tree was big.'

e. *Example 5*:



f. *Example 6*:



Sentence for examples 5–6:
Onazi san-biki-no neko-ga kaku kussyon ni neta.
same three-CL-GEN cat-NOM each cushion on slept
'The same three cats slept on each cushion.'

The pictures used for examples 1–4 in (172) are the same as those used in the examples for Phase 1. However, the sentences have been amended so that they are closer in structure and vocabulary to the sentences used in the actual test. Specifically, verbs are in the past tense, and a numerically quantified noun is used in examples 1 and 2, instead of proper nouns (i.e., *Hutari-no otoko* 'two men' instead of *Sumisu-san to Buraun-san* 'Mr Smith and Mr Brown'). Examples 5 and 6 were new for Phase 3, and were included to draw the participants' attention to the possible complexity of the pictures, and to highlight the need for careful scrutiny of the details before rating the picture-sentence match. Attention was drawn to the clocks in Example 5 indicating that a sequence of events took place; and to the colours and patterns of the cats in both examples. In Example 5 the cats' markings show that they are the same three cats, but in Example 6 the colours and markings indicate nine different cats. As noted in Section 4.1.3.2, the oral explanation also included the instruction to think of all pictures as depicting events that happened 'yesterday', so as to make the use of the past tense in the test more plausible.

The oral explanation included time for the participants to judge each example one by one, with discussion of the judgements and a chance to ask questions. This was intended to familiarise the participants with the judgement scale as well as with the format of the test and the types of pictures and sentences that might occur.

The total time taken to administer the test, including time for the questionnaires and instructions, was about 45 minutes.

4.2.4. Results

In analysing the results, responses of '+2' and '+1' on the rating scale were considered to indicate acceptance of a picture-sentence pairing, and responses of '–2' and '–1' were considered to indicate rejection. Responses of '0' show neither acceptance nor rejection. It is unclear whether ratings of '0' show that the test item is neither acceptable nor unacceptable, or whether they show that the participant did not know how to judge the test item.[126] This could be problematic if the rate of selection of '0' were very high. However, it is negligibly small in the present data set: only 3.19% of all responses were '0' (2.79% for the native Japanese group,

[126] This is an inherent problem of scales with a mid-point category. See Sorace (1996) for discussion. The mid-point category was removed in the main study (see Chapter 5).

3.02% for the learner group and 3.53% for the native English group), and these were generally evenly distributed throughout the data set. Thus the ambiguity of '0'-responses is not considered a problem in the interpretation of the results. In addition, there were very few instances of missing or illegible responses: just two out of 1150 responses in the native Japanese group, one out of 828 in the learner group, and one out of 1334 in the native English group.

4.2.4.1. Results for distractor items

The native Japanese and native English responses to the distractor items were examined first, in order to determine (i) whether the distractor items were reliable, with low rates of inter-participant variation confirming that these items were unambiguous; (ii) whether data from any participants should be excluded from the analysis due to a high number of incorrect ratings on distractor items. The average ratings and standard deviations on the distractor items by the native English and native Japanese groups are shown in Table 16.

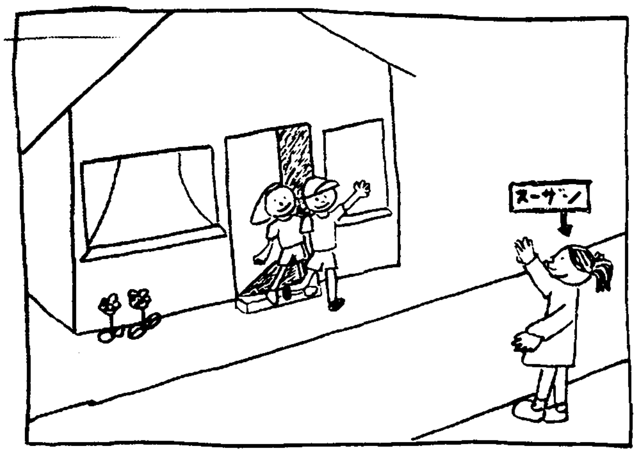
Table 16: Mean ratings and standard deviations for the distractor items in native Japanese (JJ) and native English (EE)

item	JJ (n=33) mean (SD)	EE (n=29) mean (SD)
D01	−1.76 (0.44)	−1.55 (1.09)
D02	1.30 (0.92)	1.59 (0.57)
D03	1.67 (0.69)	1.90 (0.41)
D04	−0.12 (1.71)	0.04 (1.86)
D05	1.58 (0.94)	1.97 (0.19)
D06	−0.85 (1.58)	−1.93 (0.37)
D07	−1.85 (0.36)	−1.86 (0.74)
D08	−2.00 (0.00)	−2.00 (0.00)
D09	1.59 (0.84)	1.90 (0.41)
D10	−1.67 (0.74)	−0.90 (1.52)
D11	1.94 (0.24)	1.97 (0.19)
D12	−1.79 (0.55)	−1.90 (0.41)
D13	−1.73 (0.98)	−1.72 (0.88)
D14	−1.76 (0.97)	−1.79 (0.77)

Items with a mean rating of between 1.00 and −1.00 and a standard deviation greater than 1 were considered to be potentially unreliable. Thus, for the native Japanese group, D04 (mean = −0.12, SD =1.71) and D06 (mean = −0.85, SD=1.58) were potentially unreliable, and for the native English group, D04 (mean = 0.04, SD=1.86) and D10 (–0.90, 1.52). Examination of these items confirmed that they were problematic, as detailed below.

Item D04, the distractor that produced a high level of variation in the responses of both native groups, is given in (173):

173. *Distractor item D04*:



(Translation of word in box: Susan)

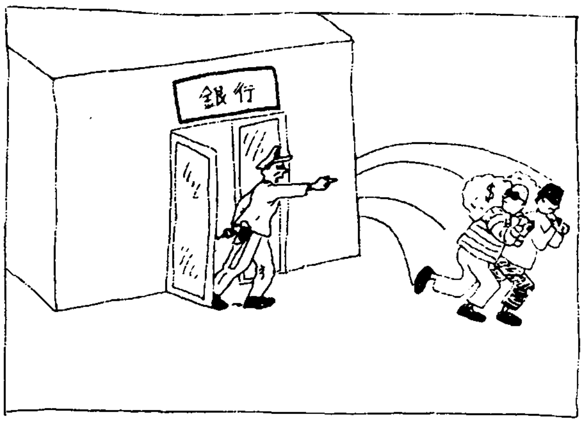
Kodomo-ga hutari-tomo Suuzan ni te-o hutta
child-NOM both Susan at hand-acc waved
'Both children waved at Susan.'

It was expected that participants would judge the picture and sentence for D04 as a mismatch (i.e., they would select '–2' or '–1') because the child on the left is not waving. However, on critical re-examination of the picture, the evident confusion was probably to the fact that the girl's hand is partially raised and she is looking towards Susan, so participants may have assumed that she was about to wave.[127]

The second confusing distractor for the Japanese group is given in 174.

[127] In the post-test questionnaire, one native English participant actually commented that it had been unclear whether or not the girl was waving in this picture.

174. *Distractor item D06*:



Translation (above door): Bank

Keisatukan-ga hutari-tomo dorobou-o mita
police officer-NOM both.HUMAN thief-ACC saw
'Both policemen saw the thieves.'

It was expected that this picture-sentence pairing would be judged unacceptable. However, eight of the 33 native Japanese participants judged the pairing acceptable, selecting +2 or +1, and one participant selected 0. Further investigation of the Japanese sentence in (174) reveals that it is, after all, genuinely ambiguous. The interpretation in which *hutari-tomo* 'both.HUMAN' modifies *dorobou* 'thief' can be argued, following Terada (1990) and Kawashima (1998), to derive by scrambling from the unambiguous structure in (175).[128]

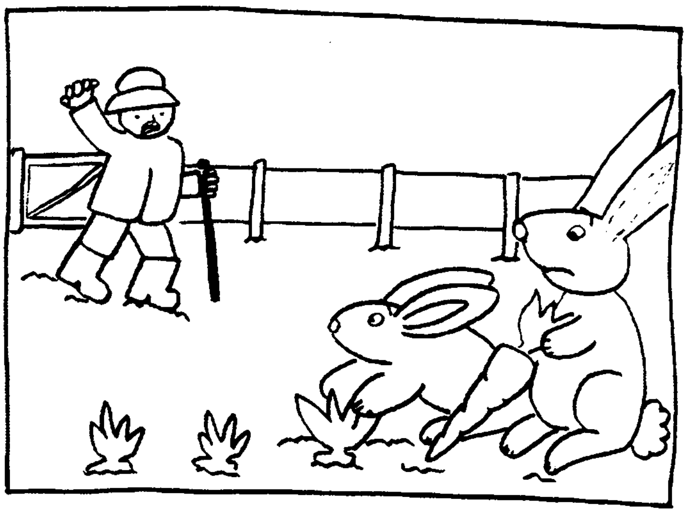
175. Keisatukan-ga [dorobou-o hutari-tomo] mita.
police officer-NOM thief-acc both.HUMAN saw
'The police officer(s) saw both thieves.'

The distractor item is therefore unreliable.

Finally, item D10, the remaining potentially unreliable distractor for the native English speakers, is given in 176:

[128] Thanks to Kyoko Oga for discussion of this point.

176. *Distractor item D10*:



Usagi-ga ni-hiki-tomo ninzin-o nusunda.
rabbit-NOM both.ANIMAL carrot-ACC stole
'Both rabbits stole the carrots.'

It was expected that the picture-sentence pair would be judged acceptable, and the item would receive ratings of +2 or +1. Indeed, for the native Japanese group, this is what happened. However, only eight of the 29 native English participants selected +2 or +1. One selected 0, and the remaining 20 selected –1 or –2. Perhaps the use of the definite article *the* in *the carrots* could be behind the difference between the native Japanese and the native English judgements. Four carrots are depicted, so if *the carrots* is interpreted to mean 'all the carrots in the picture', then a 'mismatch' rating of –1 or –2 is motivated, since the picture indicates that only two carrots were stolen. The noun *ninzin* 'carrot' in the Japanese sentence is not specified for definiteness so a similar problem does not arise.

It was decided that the results for these distractor items (D04 for both the native English and native Japanese groups, D06 for the native Japanese group and D10 for the native English group) should be ignored when considering whether to exclude any participants' data due to 'wrong' answers on the distractors. Thus, for each native group, 12 distractor items remained. The criterion for inclusion in the analysis was set at 11 or more 'correct' responses to the remaining 12 items. To repeat, on matching picture-sentence pairs, ratings of +2 or +1 were considered correct; and on mismatched picture-sentence pairs, ratings of –2 and –1 were considered correct. Ratings of 0 were never considered correct. On these grounds, the results of eight native Japanese participants were excluded because they had fewer that 11 correct distractor responses. The results of 25 native Japanese participants remained for analysis. No native English participants were excluded.

The distractor results for the learners of Japanese are presented in Table 17.

Table 17: Mean ratings and standard deviations for the distractor items by learners of Japanese (EJ)

item	EJ (n=18) mean (SD)
D01	−1.89 (0.47)
D02	1.56 (0.78)
D03	2.00 (0.00)
D04	−0.61 (1.61)
D05	1.94 (0.24)
D06	−0.17 (1.86)
D07	−2.00 (0.00)
D08	−2.00 (0.00)
D09	1.94 (0.24)
D10	−1.72 (0.69)
D11	2.00 (0.00)
D12	−1.28 (1.32)
D13	−1.56 (1.15)
D14	−1.72 (0.96)

It is clear from Table 17 that the learners' responses to items D04 and D06 were comparable with the native Japanese responses: the mean ratings on these two items are between 1 and −1 and the standard deviations are greater than 1. Since these items were deemed unreliable for native Japanese speakers, they were also ignored for the learners. Excluding responses to these two items, none of the learners had more than one 'wrong' response to the distractors, so all 18 remained in the analysis.

4.2.4.2. Results for the test items

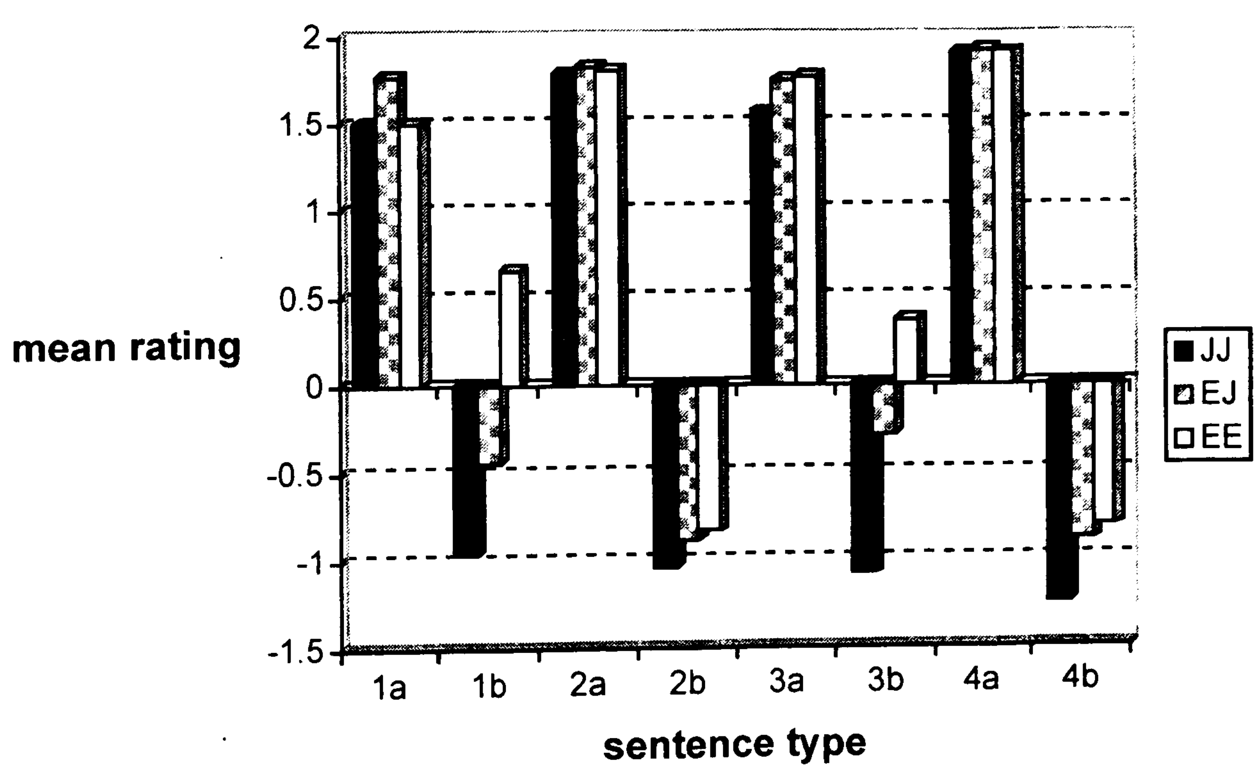
The mean ratings and standard deviations on all eight test types are presented in Table 18.

Table 18: Mean ratings on the Phase 2 QP-QP sentences in native Japanese (JJ), L2 Japanese/L1 English (EJ) and native English (EE) (Scale: –2 = 'unacceptable'; 2 = 'perfectly acceptable')

type	subject QP/object QP; scope	JJ (n=25)		EJ (n=18)		EE (n=29)	
		mean	SD	mean	SD	mean	SD
1a	*dareka* 'someone'/*dono N-mo* 'every N'; S>O	1.51	0.36	1.78	0.33	1.50	0.51
1b	*dareka* 'someone'/*dono N-mo* 'every N'; O>S	–0.97	0.74	–0.44	1.32	0.66	1.21
2a	*dareka* 'someone'/*subete-no N* 'all the N'; S>O	1.80	0.36	1.83	0.41	1.81	0.32
2b	*dareka* 'someone'/*subete-no N* 'all the N'; O>S	–1.04	0.87	–0.89	0.94	–0.84	1.17
3a	NumP/*dono N-mo* 'every N'; S>O	1.57	0.48	1.75	0.40	1.78	0.31
3b	NumP/*dono N-mo* 'every N'; O>S	–1.09	0.96	–0.29	1.26	0.37	1.34
4a	NumP/*subete-no N* 'all the N'; S>O	1.91	0.23	1.93	0.19	1.92	0.22
4b	NumP/*subete-no N* 'all the N'; O>S	–1.25	0.99	–0.96	0.81	–0.82	1.02

Figure 1 shows the mean ratings for each group in bar-chart form.

Figure 1: Mean ratings on a scale of –2 to 2 for the Phase 2A QP-QP sentences types, by group



A repeated measures ANOVA (*subject QP x object QP x scope x group*) on the data yields no significant main effect or interaction effect for the subject QP variable. The other variables—object QP, scope, and group—yield significant main effects and interaction effects, as follows: group, $F(2,69) = 7.78$, $p<.01$; object QP,

$F(1,69) = 35.9, p<.001$; scope, ($F(1,69) = 473.96, p<.001$); object QP x group, $F(2,69) = 29.42, p<.001$; scope x group, $F(2,69) = 6.32, p<.01$; object QP x scope, $F(1,69) = 79.32, p<.001$; and object QP x scope x group, $F(2,69) = 13.96, p<.001$.[129] The ANOVA results are given in full in Appendix 2C.

The sources of the significant ANOVA results are generally clear from Table 18 and Figure 1. There is an obvious effect of scope across all groups, with linear scope types (the 'a' types) receiving ratings of 1.5 or higher, while inverse scope types (the 'b' types) receive much lower ratings, ranging from –1.25 (JJ on Type 4b) to 0.66 (EE on Type 1b). Considering next the 'group' variable, it is clear that the native English group behaves differently from the other two groups on Types 1b and 3b (object QP = *every*), with positive mean ratings (0.66 and 0.37, respectively), while the other two groups have negative mean ratings for the same types. Post hoc Games Howell tests reveal that, on Type 1b, the native English ratings differ significantly from the native Japanese ($p<.001$) and from the L2 Japanese ($p<.05$) ratings. On Type 3b, there is again a significant difference between native English and native Japanese ($p<.001$), but not between native English and L2 Japanese ($p=.22$). These differences between the English group on the one hand and the native Japanese and L2 Japanese on the other account for the significant main effect for group. The significant main effect of the object QP variable and all the significant interactions must also derive from native English acceptance of Types 1b and 3b compared with their rejection of Types 2b and 4b (object QP = *all*; mean EE ratings, –0.84 and –0.96, respectively). These four inverse scope types are differentiated by their object QP: *every* for Types 1b and 3b, *all* for Types 2b and 4b. The fact that, in contrast to the native and L2 Japanese groups, the native English group accepts inverse scope when the object QP is *every* but not when it is *all* must account for the significant effect of the object QP variable and the significant interactions among the object QP, scope and group variables.

To summarise the overall pattern of the results, all three groups rate linear scope (Types 1a, 2a, 3a and 4a) highly acceptable (mean ratings ≥1.5). By contrast, inverse scope (Types 1b, 2b, 3b and 4b) is rejected by the native Japanese (mean

[129] Levene's test for homogeneity yielded significant results ($p>.05$) for Types 1a, 1b and 3a. Thus homogeneity of variance cannot be assumed across all the data. Consequently, F-ratios and p-values may not be as reliable as if homogeneity of variance were assured. Games-Howell tests are used for post hoc analyses, since this procedure is more accurate when homogeneity of variance is not assumed.

ratings ≤–0.97) and L2 Japanese (mean ratings ≤–0.29) groups, and also by the native English group when the object QP is *all* (mean ratings ≤–0.82 on Types 2b and 4b). However, when the object QP is *every*, the native English group accepts inverse scope (mean ratings ≥0.37 on Types 1b and 3b).

In addition to this overall picture, two further details are important. First, all three groups exhibit greater variation (as measured by the standard deviations in Table 18) in their responses to inverse scope test items (the 'b' types) than to linear scope test items (the 'a' types). On the linear scope items, standard deviations are 0.51 (EE on Type 1a) or lower, whereas on the inverse scope items, standard deviations range between 0.74 (JJ on Type 1b) and 1.34 (EE on Type 3b). Correspondingly, the mean ratings for the linear scope types are high (≥1.5) on the scale of –2 to +2 for all three groups. However, the mean ratings for the inverse scope items are not as extreme: when inverse scope is rejected, the mean ratings are closer to –1 than –2 on the rating scale (range: –0.29 [EJ, Type 3b] to –1.25 [JJ, Type 4b]); and when inverse scope is accepted (only Types 1b and 3b in English), the mean ratings are between 0 and 1 (0.66 for Type 1b, 0.37 for Type 3b).

The second important detail concerns the L2 group. Although this group tends to reject inverse scope, the mean ratings on the inverse scope items when the object QP is *dono N-mo* 'every N' (Types 1b and 3b) are not as low as the corresponding ratings by the native Japanese group: EJ –0.44 compared with JJ –0.97 on Type 1b; EJ –0.29 compared with JJ –1.09 on Type 3b. The difference between the two groups is close to significant on Type 3b (p=.078 on post hoc Games Howell test), although not on Type 1b, (p=.3). The implications of this and the other findings reported in this section are the subject of the following discussion.

4.2.5. Discussion

The goals of this pilot study were: (i) to find out whether the results of the preliminary test (Phase 1) with native speakers of Japanese would be replicated, thereby confirming the reliability of the test; (ii) to find out whether native English speakers would indeed accept inverse scope on an English version of the test, thereby further confirming the sensitivity of the test instrument; (iii) to trial the test instrument with English-speaking learners of Japanese and test the predictions in (166a–b), repeated here as (177a–b):

177. a. *Prediction 1: (L1 Transfer)*
Due to L1 transfer, learners whose L1 is English will allow non-target-like ambiguity in Japanese doubly-quantified sentences.

b. *Prediction 2: (L2 Access to UG)*
Due to L2 access to UG, English-speaking learners of Japanese *can* acquire Japanese scope rigidity (despite underdetermination by the evidence).

The first section of this discussion (Section 4.2.5.1) focuses on the native data and goals (i) and (ii), the second (Section 4.2.5.2) on the learner data and goal (iii).

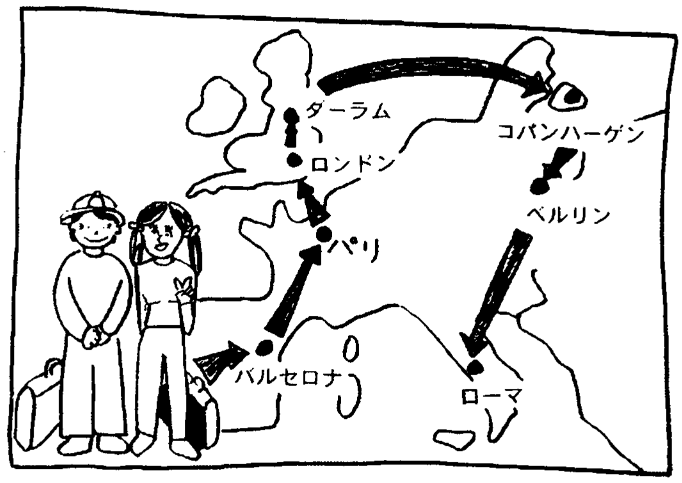
4.2.5.1. Discussion of the native Japanese and native English data, Phase 2A

The native Japanese and native English data seem to confirm the claims of the theoretical literature: inverse scope interpretations of the doubly quantified sentences in the test were generally unacceptable in native Japanese, but acceptable in native English when the object quantifier was *every*, although when the object quantifier was *all*, inverse scope was rejected in English. For Japanese, this corroborates the findings of Phase 1: Types 1a–b, 2a–b, 3a–b and 4a–b in Phase 1 were the same as Types 1a–4b in Phase 2A (although the tokens were different) and yielded the same pattern of acceptance of the linear scope items and rejection of the inverse scope items. The results of Phase 1 are thus replicated, indicating that the test instrument is reliable. For native English, the response pattern corroborates the findings of Lee *et al.* (1999a, 1999b), using a different test instrument, as discussed in Chapter 3. This, along with the fact that the acceptance/rejection pattern of the native English results is as predicted according to the theoretical literature, further indicates that the test instrument is reliable.

The finding that there was more variation in responses to inverse scope test items than linear scope test items, and that mean ratings on these items were closer to the middle of the rating scale than the extremes of the rating scale, also replicates what was found in Phase 1 for native Japanese. The theoretical literature predicts that inverse scope should be rejected outright in native Japanese (i.e., with ratings close to –2 on the scale used in the present experiment), and that it should be completely acceptable (i.e., mean ratings close to +2) in native English when the object quantifier is *every*. Clearly, such outright rejection or acceptance of inverse scope did not occur in the test. In the discussion of Phase 1 (Section 4.1.3), the possibility was raised that the lack of outright rejection of inverse scope in Japanese might be due to aspects of the experimental design. Consequently, a number of refinements were made to the test instrument for Phase 2, in order to increase reliability and potentially lower variation. Since the Phase 2 results nonetheless yield a similar degree of variation, it seems that, with this test instrument, interpretation of inverse scope entails a certain amount of variation while still generally confirming the claims of the theoretical literature. As noted in Chapter 2, it is frequently observed that interpretation of quantifier interactions is not easy: judgements are often vague and variation among individuals is common. In addition, the experimental study of Kurtzman & MacDonald (1993) indicated that linear scope interpretations of doubly-quantified sentences are more readily available than inverse scope interpretations (see Chapter 3). The variation in judgements on the inverse scope types in the present study thus may serve to quantify the reality that inverse scope contexts are hard to interpret. However, the possibility that the variation may be an artefact of the test instrument is not ruled out. Specifically, the picture contexts may lead to individual variation. For example, the picture used for item 11a (illustrated in (178) below) was intended to show that the two tourists visited the seven cities on the map. However, some individuals may think that, because the two tourists are on the left of the picture and the arrows point rightwards, the tourists are at the beginning of their trip, and the picture does not indicate whether or not they actually completed the trip.[130] Therefore, the picture-sentence match could be judged unacceptable on these grounds, and not because the participant rules out the linear scope interpretation that this picture is intended to illustrate.

[130] This criticism of item 11a was raised at a work-in-progress discussion of the study at the University of Durham, June 2002.

178. *Item 11a:*



Translation (following arrows): Barcelona, Paris, London, Durham, Copenhagen, Berlin, Rome

Hutari-no kankoukyaku-ga dono mati-mo kenbutu-sita.
two.HUMAN-GEN tourist-NOM every town-QPt visited
'Two tourists visited every city.'

A solution to such problems could be to use precise, written contexts instead of pictures. Thus, an additional small-scale pilot test was conducted, to investigate whether written contexts might yield lower levels of variation than picture contexts. This test, Phase 2B, is detailed in Section 4.3.

4.2.5.2. Discussion of the L2 Japanese data, Phase 2A

Moving on to the third goal of the present pilot test, the results for the English learners of Japanese are compatible with the two predictions in (177), namely that (177a) learners whose L1 is English allow non-target-like ambiguity in Japanese doubly-quantified sentences due to influence from their L1; and (177b) English-speaking learners of Japanese can nonetheless acquire Japanese scope rigidity. The key test types relevant to the two predictions are those with the object QP *dono N-mo* 'every N' and inverse scope: Types 1b (subject = *dareka* 'someone') and 3b (subject = NumP). On the English version of these types, the inverse scope interpretation is acceptable, (mean ratings: 0.66 on Type 1b, 0.37 on Type 3b) while it is unacceptable in native Japanese (–0.97 on Type 1b, –1.09 on Type 3b). The learners of Japanese rejected inverse scope on these two types (mean ratings: –0.44 on Type 1b, –0.29 on Type 3b), but their rates of rejection were not as high as those of the native Japanese group. In effect, their mean ratings on Types 1b and 3b fell between the low native Japanese ratings and the high native English ratings. This could be the result of the interlanguage grammar of some of the learners still being influenced by the L1 and consequently allowing inverse scope while the grammar of other learners in the group had already undergone reconstruction to rule out inverse scope. Examination of the consistency with which individuals rejected or accepted inverse scope is informative here. Table 19 compares learners of Japanese with the native English and Japanese control groups, with respect to individual consistency of response. 'Consistent rejection' of a type is defined as ratings of '–1' or '–2' on at least three of the four tokens for that type; 'consistent acceptance', as ratings of '+1' or '+2' on at least three of the four tokens for that type; and 'inconsistency', as fewer than three positive or three negative ratings on the four tokens for that type. (No individual in either group consistently selected '0' for any type.) Six response categories thus emerge, as shown in Table 19.

Table 19: Consistency of individuals on Types 1b and 3b

Response categories (Type 1b = dareka-ga dono N-mo V.PAST 'someone V.PAST every N.') (Type 3b = NumP-ga dono N-mo V.PAST 'NumP V.PAST every N.')	**No. (%) of individuals** EJ (n=18)	EE (n=29)	JJ (n=25)
1. consistent rejection on 1b and 3b	6 (33.3)	4 (13.8)	16 (64.0)
2. consistent acceptance on 1b and 3b	3 (16.7)	13 (44.8)	1 (4.0)
3. inconsistency on 1b and 3b	1 (5.6)	2 (6.9)	3 (12.0)
4. consistent rejection on 1b or 3b, inconsistency on the other	2 (11.1)	1 (3.5)	4 (16.0)
5. consistent acceptance on 1b or 3b, inconsistency on the other	2 (11.1)	3 (10.4)	1 (4.0)
6. consistent acceptance on 1b or 3b, consistent rejection on the other	4 (22.2)	6 (20.7)	0

The native Japanese data in Table 19 show that 16 of the 25 participants consistently rejected object-wide scope on both of the relevant test types, and a further four consistently rejected object-wide scope on one of the test types and were inconsistent on the other. These data contrast clearly with the data for the native English controls, which give a picture of what might transfer from the L1 on an individual level. The largest grouping is, as expected, in Category 2: 13 of the 29 native English controls consistently accepted inverse scope. However, four out of 29 fall into Category 1: they consistently rejected inverse scope even though it is theoretically available in English. The remaining native English controls fall into Categories 3–6, indicating some degree of inconsistency in their responses. This shows that, on the present test instrument, some English-speaking learners of Japanese might reject inverse scope in Japanese because that is what they would do in English (if they are like the four participants in Category 1), and not because their grammar has been restructured to resemble the native Japanese grammar. However, the proportion of participants in Category 1 (rejection of inverse scope) is considerably smaller in the native English group (13.8%) than in the L2 Japanese group (33.3%). Similarly, the proportion in Category 2 (acceptance of inverse scope) is considerably larger in the native English group (44.8%) than in the L2 Japanese group (16.7%). In other words, the L2 Japanese data do not look exactly like the native English data on an individual level. It is possible that the interlanguage of each of the six learners of Japanese in Category 1 has indeed undergone restructuring from a grammar that accepted inverse scope to one which rejects it; while that of the each of the three learners in Category 2 has not yet undergone restructuring, and thus these individuals consistently accept inverse scope, as in English. Further evidence to support such a conclusion would come from proficiency data. If the three L2 learners in Category 2 are of demonstrably lower L2 Japanese proficiency than the six in Category 1, this would support the theory that, at the initial (or closer-to-initial) state of L1 acquisition, the learners' interlanguage grammar is like their L1, while at more advanced levels it becomes more like the target language grammar. This highlights the need for proficiency data, which were not available in Phase 2, but which were collected in the main study, Phase 3.

To summarise with respect to the predictions in (177), the overall data and the individual-level data are compatible with the predictions. The fact that the learners do not reject inverse scope to the same degree as native Japanese speakers suggests that they are influenced by the acceptability of inverse scope (when the object quantifier is *every*) in their L1. However, the fact that at least some learners consistently reject inverse scope in Japanese suggests that their interlanguage grammar has undergone restructuring with respect to quantifier interpretation. L2 Japanese proficiency data would enable clearer conclusions to be drawn.

4.3. Phase 2B: A further pilot study with native speakers of Japanese

Phase 2B was a small-scale pilot study to investigate whether written contexts instead of picture contexts for the doubly-quantified sentences would yield lower levels of variation in the responses to inverse scope items.

4.3.1. Design, participants and procedure

The design differed from that of Phase 2A in that the object QP was not varied: only *dono N-mo* ‘every N’ was used. The test sentences were exactly the same as those used for Types 1a–b and 3a–b in Phase 2A. The design is summarised in Table 20. (Types 1a–b and 3a–b in Table 20 correspond exactly to 1a–b and 3a–b in Phase 2A. The labels ‘2a’ and ‘2b’ are not used in this experiment.)

Table 20: Test types for Phase 2B

type	subject QP	scope	example
1a	*dareka* ‘someone’	S>O	Dareka-ga dono neko-mo nadeta. someone-NOM every cat-QPt stroked ‘Someone stroked every cat.’
1b		O>S	
3a	NumP	S>O	Sannin-no onnanoko-ga dono tako-mo ageta. three-GEN girl-NOM every kite-QPt flew ‘Three girls flew every kite.’
3b		O>S	

The scope variable was manipulated by means of written contexts, and each test sentence was linked to its written context using the word *tumari* ‘in other words’. Examples for Types 1a and 1b are given in (179). (Only the actual Japanese text and its English translation are given for the contexts in (179), since a word-for-word gloss is not relevant here. Note that the QPs used in the test sentences [*dareka* ‘someone’, NumP, and *dono N-mo* ‘every N’] did not appear in the written contexts. In the actual test, the participants saw only the Japanese text with no English translation.)

179. a. *Item 3a (Type 1a):*

Context:
朝日農場には猫が何匹かいる。昨日、この農場に小学生のグループが見学に来た。一人の小学生は、猫をみんな集めて、なでた。
‘There are several cats on Asahi Farm. Yesterday, a group of primary school children went on a school trip to the farm. One of the children gathered all the cats together and stroked them.’

Test sentence:
つまり、だれかがどの猫もなでた。
Tumari, dareka-ga dono neko-mo nadeta.
in other words, someone-NOM every cat-QPt stroked
‘In other words, someone stroked every cat.’

b. *Item 3b (Type 1b):*

Context:
校門の近くに五匹の猫が入った箱が捨てられていた。五人の子供がその捨て猫を見つけた。子供は一人一人、違う猫を箱から出して、なでた。
'Five cats were abandoned in a box outside the school gates. A group of five children found them. Each child picked up a different cat from the box and stroked it.'

Test sentence:
(as for Item 3a in (179a) above)

There were four tokens of each type, and these were mixed randomly with scrambled variants of each token and with 12 distractor items.[131] The complete test set is given in Appendix 2D.

Ten native speakers of Japanese participated in the experiment. Their average age was 21 (range: 19–23). All were university students. Nine had arrived in the UK less than three weeks prior to participating in the experiment, for an intensive English course; one had been in the UK for two years. None of the Phase 2B participants had participated in Phase 2A.

The procedure was as for Phase 2A. The written contexts were projected onto a screen for 15 seconds each, then the test sentence was projected below the written context so that the two could be viewed together for a further 10 seconds. Both contexts and test sentences were also presented orally by the researcher (a near-native speaker of Japanese). The participants rated how well each sentence matched the context using the same five-point scale as in Phase 2A (–2 = 'they don't match at all'; +2 = 'they match perfectly').

4.3.2. Results and discussion

As in Phase 2A, ratings of –2 and –1 are considered to indicate rejection of a context-sentence match, and ratings +2 and +1 are considered to indicate acceptance. Ratings of '0' indicate neither acceptance nor rejection. The rate of selection of 0

[131] Word order—canonical v. scrambled—was a further variable in Phase 2B. However, this section focuses on discovering whether there was any difference in levels of variation in ratings between the picture context experiment (Phase 2A) and the written context experiment. Phase 2A did not include scrambled test items. Thus, the results for the scrambled sentences in Phase 2B are not discussed. It is acknowledged that the use of scrambled test sentences and different distractors in Phase 2B constitutes an extraneous variable that could have affected the participants' responses.

was very low: 24 instances out of 440, spread evenly across the data. There were two missing or illegible responses.

No participant gave more than one unexpected response to any of the distractor items, so the data from all participants were included in the analysis. Table 21 compares the mean ratings and standard deviations for Types 1a, 1b, 3a and 3b in the present experiment with those of the native Japanese group in Phase 2A.

Table 21: Mean ratings on Phase 2A (picture contexts) and Phase 2B (written contexts) Types 1a, 1b, 3a and 3b in native Japanese (Scale: –2 = 'unacceptable'; 2 = 'perfectly acceptable')

type	subject QP/object QP; scope	JJ Phase 2A (n=25)		JJ Phase 2B (n=10)	
		mean	SD	mean	SD
1a	*dareka* 'someone'/*dono N-mo* 'every N'; S>O	1.51	0.36	0.08	1.20
1b	*dareka* 'someone'/*dono N-mo* 'every N'; O>S	–0.97	0.74	0.13	1.07
3a	NumP/*dono N-mo* 'every N'; S>O	1.57	0.48	1.33	0.33
3b	NumP/*dono N-mo* 'every N'; O>S	–1.09	0.96	–0.92	0.84

A repeated measures ANOVA run on the native Japanese data of both Phases 2A and 2B yields a significant main effect for scope, $F(1,33)=137.06, p<.00$, but no significant main effect for subject QP or group (i.e., the Phase 2A picture context group v. the Phase 2B written context group). All the interactions with scope are significant: scope x group, $F(1,33)=22.31, p<.001$; subject QP x scope, $F(1,33)=45.05, p<.001$; subject QP x scope x group, $F(1,33)=32.73, p<.001$. The interaction between subject QP and group is not significant. See Appendix 2E for full statistical details.

From Table 21 it is clear that the Phase 2B results for Types 3a and 3b were very similar to those for Phase 2A: linear scope (Type 3a) is accepted with a mean rating of 1.33 in Phase 2B compared with 1.57 in Phase 2A; and inverse scope (Type 3b) is rejected with a mean rating of –0.92 in Phase 2B compared with –1.09 in Phase 2A. Independent samples *t*-tests confirm that there is no significant difference between Phase 2A and Phase 2B on Type 3a ($t(33)=1.48, p=.15$) and Type 3b ($t(33)=-.5, p=.62$). However, the results for Types 1a and 1b are different: while linear scope (Type 1a) is clearly accepted in Phase 2A (mean rating = 1.51) and inverse scope (Type 1b) is rejected (mean rating = –0.97), in Phase 2B there is very little difference between the Type 1a and Type 1b mean ratings (0.08 on Type 1a, 0.13 on Type 1b), and both ratings are very close to the middle of the scale. Correspondingly, the Phase 2B standard deviations on these two types are relatively high (≥1.07). Thus, there was more individual variation on these types in Phase 2B, with the written contexts, than in Phase 2A, with the picture contexts. Independent samples *t*-tests show significant differences between Phase 2A and Phase 2B for both Types 3a ($t(9.64)=3.71, p<.05$ (equal variances not assumed)) and 3b ($t(33)=-3.49, p<.05$).

Examination of individual test items in Phase 2B shows that the level of variation is approximately equal across the four Type 1a and four Type 1b tokens, with standard deviations ranging from 1.45 (item 1b) to 1.76 (item 2a). Thus the source of the variation in Phase 2B does not derive from any individual token or tokens producing a much higher level of variation than the others. The results thus indicate that, on Types 1a and 1b, the use of written contexts (Phase 2B) was more confusing (i.e., led to greater inter-participant variation) than the use of picture contexts (Phase 2A).

A post-test questionnaire, along with informal post-test interviews with some of the participants, revealed two possible reasons for the evident confusion in Phase 2B: (i) some difficulty was experienced in identifying the word *dareka*/'someone' with a person who had been specified in the scenario; (ii) the complexity of the scenarios required 'puzzling over', and some lapses occurred in the attentiveness to reading. The first reason affects Types 1a and 1b specifically (since these types had the subject *dareka* 'someone'), while the second reason applies to all the test sentences in Phase 2B.

The first reason can be exemplified with respect to item 2a shown in (180).

180. *Phase 2B, Item 2a:*

> 先週、ある男が街角にある喫茶店で、メニューにある料理を全部注文した。その男は一番おいしい料理がどれか知りたいのだと言って、全部の料理を食べてみた。
> 'Last week, a man ordered one of everything on the menu at the corner café. He said he wanted to see which dish was tastiest, and he tried each one.'
>
> つまり、だれかがどの料理も食べてみた。
> Tumari, dareka-ga dono ryouri-mo tabete-mita.
> in other words, someone-NOM every dish-QPt eat-tried
> 'In other words, someone tried every dish.'

Some participants recounted feeling that *dareka* 'someone' in the test sentence could not refer to ある男 *aru otoko* 'a (particular) man' in the scenario, because they felt *dareka* should indicate an unspecified person. Thus, five of the 10 Phase 2B participants gave this scenario-sentence pair a rating of –2 or –1, even though it is a linear scope item and the expected rating was +2. By contrast, in Phase 2A all 25 of the native Japanese participants gave the picture-sentence pair for item 2a a positive rating. Clearly, there was no problem allowing *dareka* to refer to the man in the picture.

The second possible reason for higher levels of variation in Phase 2B is exemplified with the scenario for item 3b, presented in (179b), above. The romanisation and gloss of the final sentence of the scenario is presented in (181).

181. Kodomo-wa hitori-hitori, tigau neko-o hako kara dasite, nadeta
child-TOP each.HUMAN, different cat-ACC box from take out stroked
'Each child took a different cat out of the box and stroked it.'

One participant explained that she had misapprehended *hitori-hitori* 'each.HUMAN' as *ippiki-ippiki* 'each.ANIMAL'. This error makes the floating quantifier refer to the cats, and clouds the crucial fact that each individual child had taken out one cat. Although probably a one-off error, this may be indicative of the types of problems behind the participants' frequent comments that the scenarios had been confusing, or that sometimes there had not been enough time to work out what was happening. Comments about confusing pictures were rare in the post-task questionnaire for Phase 2A, despite asking specifically whether any of the pictures had been confusing. In short, it appears that reading/hearing a written scenario was a much more confusing and onerous task than looking at a picture scenario. Moreover, in the case of Types 1a and 1b, confusion—and hence variation in responses—was exacerbated by the difficulty, described above, of allowing *dareka* 'someone' to refer to a person specified in the context.

A crucial problem with written contexts for the main study of this research project has already been discussed (see Chapter 3). Namely, since the main test must be rendered into four different languages (Japanese, English, Chinese, and Korean), assuring consistency of meaning across the four languages is a serious challenge. If the use of written contexts in this pilot study had yielded considerably less variation than the picture contexts in Phase 2A, it could have been worth attempting to overcome the translation problem and using written contexts in the main study. However, since written contexts appeared more difficult to process than pictures, it was decided to continue to use picture contexts for the main study. In an attempt to further decrease possible variation due to idiosyncratic interpretations of the pictures, the pictures were critically re-examined for the main study and redrawn or altered where necessary.

4.4. Conclusion

This chapter has detailed the development and pilot testing of an acceptability judgement task to investigate scope interpretation in L2 Japanese. Phase 1 of the development trialled a picture-based acceptability judgement task with native speakers of Japanese. The task yielded the expected result: participants generally rejected inverse scope in doubly-quantified Japanese sentences. The instrument was refined and improved for a full-scale pilot study in Phase 2, with three participant groups: native Japanese, native English, and L1 English-L2 Japanese. The native Japanese and native English results were as expected based on the theoretical literature: inverse scope was again generally rejected in Japanese (confirming the reliability of the test instrument), while it was accepted in English when the object quantifier was *every*, but not *all*. The L2 Japanese group tended to reject inverse scope, although this group's rate of rejection was not as high as in the native Japanese group. This suggested some effect of influence from their L1 (English) in which inverse scope is permitted. However, it was observed that an objective measure of the learners' L2 Japanese proficiency is crucial for drawing conclusions about the L2 data. Finally, a further small-scale pilot study was described, in which written contexts were provided for the QP-QP sentences instead of picture contexts. This test was completed by native speakers of Japanese. The results showed that the method led to more variation in responses than the picture-based method. Thus, it was decided to continue to use the picture-based method for the main study, but to further improve the pictures where possible. A number of other changes were also made to the design for the main study, with the aim of further increasing the reliability of the test instrument. These are reported in the following chapter.

5. The main studies

5.0. Introduction

This chapter presents Phase 3 of the experimental work conducted for this dissertation, namely two full-scale comparative interlanguage studies of (i) QP-QP scope interpretation and (ii) *Wh*-QP scope interpretation in L2 Japanese. Acceptability judgement tasks (developed through the pilot testing detailed in Chapter 4) were used to investigate these two phenomena with learners of Japanese whose L1s are Chinese, English, or Korean, in order to address the two research questions posed in Chapter 1 (based on Schwartz & Sprouse 2000):

182. Do L2ers show divergence with respect to a target language phenomenon *P* when their L1s are typologically distinct with respect to *P*?

183. When *P* represents an L2 poverty-of-the-stimulus problem, are L2ers able to overcome the problem and acquire *P*?

As detailed in Chapter 1, the question in (182) seeks to identify the role of L1 transfer in L2 acquisition. The question in (183) addresses the issue of the role of UG: if L2ers are able to acquire *P* under poverty of the stimulus, this would suggest that L2 acquisition is mediated by the innate mechanisms of UG. The experiments test the Full Transfer/Full Access model of L2 acquisition (Schwartz & Sprouse 1994, 1996; see Chapter 1), which claims that the initial state of L2 acquisition is the L1 grammar, and subsequent restructuring is governed by UG and motivated by failure to map the target language input to the interlanguage grammar.

For QP-QP scope interpretation, it was hypothesised that L1 English-speaking learners of Japanese would show a different developmental path from L1 Korean- and L1 Chinese-speaking learners of Japanese, because (as detailed in Chapter 2) English is typologically distinct from the other three languages in that it readily allows inverse scope in [S...O...] QP-QP sentences while the other three languages do not. Acquisition of the more restricted Japanese grammar (*O>S in SOV QP-QP sentences) represents a poverty-of-the-stimulus problem for L1 English-speaking learners of Japanese. Thus, to address the second research question (183), the QP-QP experiment additionally tested the hypothesis that advanced L1 English-speaking learners of Japanese would nonetheless be able to acquire native-like QP-QP scope interpretation.

For *Wh*-QP scope interpretation, the initial hypothesis with respect to L1 transfer was that L1 English- and L1 Chinese-speaking learners of Japanese would show a different developmental path from that of L1 Korean-speaking learners of Japanese, because English and Chinese are argued to differ from Japanese and Korean in allowing pair-list interpretations of questions with a QP-subject and *wh*-object.[132] In this case acquisition of the more restricted target language grammar (which lacks pair-list interpretations) is a poverty-of-the-stimulus problem for both L1 English- and L1 Chinese-speaking learners of Japanese. The *Wh*-QP experiment thus tested the hypothesis that these learners would nonetheless be able to acquire native-like *Wh*-QP scope interpretation at an advanced level of proficiency.

The hypotheses (which are spelled out more precisely in the relevant sections, following) are supported by the results. Two major findings are that (i) the developmental path of L1 English-speaking learners of Japanese differs from the paths of L1 Korean- and L1 Chinese-speaking learners of Japanese with respect to QP-QP scope interpretation; and (ii) advanced L1 English-speaking learners of Japanese show evidence of acquiring target-like QP-QP scope interpretation in Japanese, and advanced L1 English- and L1 Chinese-speaking learners of Japanese show evidence of acquiring target-like *Wh*-QP interpretation, despite poverty of the stimulus.

As well as investigating L2 Japanese, control studies are conducted of the relevant QP-QP and *Wh*-QP interpretations in native Japanese, Chinese, English and Korean. The control data are interesting in their own right since, as pointed out in Chapter 3, very little quantitative empirical data is available with respect to the phenomena under investigation, particularly for Japanese and Korean. The control data generally confirm the claims of the theoretical literature (Chapter 2). However, for *Wh*-QP scope interpretation, the claim that Korean lacks pair-list readings is not supported. This leads to reformulation of the hypotheses for *Wh*-QP interpretation in L1 Korean/-L2 Japanese interlanguage, as detailed in the relevant section. The L1 Korean-L2 Japanese data support the revised hypothesis.

The chapter is organised as follows. Section 5.1 gives details of elements of the experimental procedure common to both the QP-QP study and the *Wh*-QP study,

---

[132] As noted below, and detailed in full in Section 5.3, empirical data on *Wh*-QP interpretation in native Korean, collected at the same time as the L2 data, led to retrospective reformulation of this hypothesis for L1 Korean-L2 Japanese interlanguage.

namely the overall design, the participants, and the L2 Japanese proficiency task. Section 5.2 then presents the specific details of the procedure for the QP-QP study, along with the results and analysis; and Section 5.3 presents the same for the *Wh*-QP study. Finally, Section 5.4 briefly summarises the findings and conclusions.

5.1. Experimental procedure

5.1.1. Overall experiment design

The L2 experimentation for Phase 3 comprised four tasks: two acceptability judgement tasks investigating QP-QP interpretation ('QP-QP Task 1' and 'QP-QP Task 2'), one investigating *Wh*-QP interpretation, and a cloze test to measure L2 Japanese proficiency. These four tasks were also completed by native Japanese control participants. In addition, the QP-QP and *Wh*-QP acceptability judgement tasks were translated into Chinese, English, and Korean and completed by native speakers of those languages. Specific details of the QP-QP and *Wh*-QP tasks are given in Sections 5.2 and 5.3, respectively. This section details the overall procedure for obtaining all the necessary data.

The times allotted for each of the four tasks completed by the L2er participants and the native Japanese control participants were as follows:

184. *Time allotted for tasks in Phase 3:*

QP-QP Task 1:	30 minutes (12 minutes for explanation, 18 minutes for task)
QP-QP Task 2:	18 minutes (no explanation needed after QP-QP Task 1)
Wh-QP task:	20 minutes (10 minutes for explanation, 10 minutes for task)
Cloze test:	20–30 minutes (5 minutes for explanation, up to 25 minutes for task)

The times given in (184) all allow for completion of each component comfortably, without rushing. It was intended that the L2 participants would complete the four tasks over two sessions on separate days, each session lasting at most one hour (including time for entering the room, filling in personal details on the answer sheets, etc.). The two sessions were organised as shown in (185):

185. a. *Session 1:*
QP-QP Task 1
Wh-QP task

b. *Session 2:*
Cloze test
QP-QP Task 2

In practice, constraints such as room availability and participants' timetables did not always allow the breakdown into the two sessions in (185). In order to obtain large enough sample sizes, the experimentation was conducted at several different universities: three in the UK, five in Japan, and one in Korea. The data were then collapsed into the desired participant groups for analysis (e.g., the data from L1 Korean-speaking learners of Japanese came from participants at three universities in Japan and one in Korea). Each participating university had its own space and time restrictions within which the research had to be conducted. Consequently, while some L2 participants completed the tasks over two sessions on separate days, others did them on a single day with only a short break between sessions, and a number of the L1 Korean-speaking learners of Japanese did the three acceptability judgement tasks in a single session, with the cloze test alone being completed a week later.

All L2ers included in the research presented in this chapter completed all four components given in (184). The native control participants, on the other hand, did not generally do all four tasks. Individual control participants did either the two QP-QP tasks, or the *Wh*-QP task.[133] The native Japanese control participants who did the *Wh*-QP task also completed the cloze test. In addition, a further twelve native Japanese participants completed only the cloze test.

All the experimentation was administered by the author, except for the cloze test for 25 L1 Korean-speaking learners of Japanese, which was administered by those learners' Japanese language teacher following the author's protocol. Although the research was conducted at several different venues, with the result that environmental conditions could not be held constant, the author was satisfied that each venue (all university classrooms) was comfortable, well lit, free from distracting noise, and had properly-functioning audio-visual equipment.

The instructions on how to complete each task were presented at each test session by the author, either in English or in Japanese. Instructions were given in English to the L1 English-speaking learners of Japanese and the native English and native Chinese control participants (the groups who participated in the UK); instructions were given in Japanese to the L1 Chinese- and L1 Korean-speaking learners of Japanese and the native Japanese and native Korean control participants

133 Exceptions are all 26 of the native Korean controls, along with five native English controls and one native Chinese control. These participants all completed QP-QP Tasks 1 and 2, *and* the *Wh*-QP task.

(the groups who participated in Japan or Korea).[134] The oral instructions were based on the written instructions on the answer sheets and, as such, were the same from one session to the next.[135] However, as in the pilot studies (see Chapter 3). participants had the opportunity to ask questions during the presentation of the instructions and the examples, and individuals' questions varied from session to session.

The sentences in the Japanese tasks were written in standard Japanese script with *furigana* (phonetic superscript) above all but the most basic *kanji* (ideographs), so as to ensure that learners would not struggle with reading the script. In the Chinese versions of the tasks, the script was standard simplified Chinese characters; in the Korean versions, *hangul* (standard Korean script).

5.1.2. Participants

Data from 29 L1 English-speaking learners of Japanese ('EJ'), 17 L1 Chinese-speaking learners of Japanese ('CJ'), and 38 L1 Korean-speaking learners of Japanese ('KJ') were used in the research presented in this chapter.[136] These groups were divided into 'intermediate' and 'advanced' sub-groups on the basis of the proficiency test described in Section 5.1.3, following. Details of the resulting six learner groups are summarised in Table 22.

---

[134] The native Korean control participants were all members of a Japanese language class, and thus could understand Japanese. Care was taken to present instructions simply, to ensure learners' understanding.

[135] Test-specific details of the instructions are given in Sections 5.2 and 5.3.

[136] Data were originally collected from 30 EJ learners, 48 CJ learners and 43 KJ learners. Some of these data had to be excluded due to missing proficiency task and/or QP-QP Task 2 scripts (1 EJ learner, 8 CJ learners, 6 KJ learners). In addition, 23 CJ learners were excluded because 21 were bilingual in Chinese and Korean, and two gave their native language as Mongolian. (CJ learners who indicated that they were bilingual in two dialects of Chinese (e.g., Mandarin and Cantonese) were not excluded. See comment on Chinese dialects, following.)

Table 22: L2 Japanese participants[a]

group	no.	age		years living in Japan (y;mo)		details
		range	mean	range	mean	
EJ int	20	19–30	21	0;0–1;6	0;8	Students of Japanese at UK universities, resident in UK at the time of participation.
EJ adv	9	21–23	22	1;0–2;2	1;2	
CJ int	7	21–34	25	0;9–3;9	1;6	Students at Japanese universities, resident in Japan at the time of participation.
CJ adv	10	21–33	26	0;9–4;9	2;11	
KJ int	23	22–51	28	0;0–3;6	0;6	Students at universities in Japan or Korea. (KJ int: 3 participants resident in Japan at the time of participation, 20 in Korea; KJ adv: 10 participants resident in Japan at the time of participation, 5 in Korea.)
KJ adv	15	21–31	24	0;0–2;9	1;4	

[a] 'int' = intermediate; 'adv' = advanced

All of the L2er participants in Table 22 had learnt Japanese via classroom instruction: they were not naturalistic learners (although, as is evident from the table, many also had experience of living in Japan). All also had experience of studying foreign languages other than Japanese.[137] None of the EJ or KJ participants were bilingual from birth or early childhood. The CJ participants were predominantly native speakers of Mandarin Chinese, although some had other Chinese dialects as their native language, and some were bilingual in two dialects of Chinese.[138] All participants volunteered to take part in the study after being invited to do so by their Japanese language teachers.

As well as the L2er participants, native speakers of Japanese ('JJ'), Chinese ('CC'), English ('EE') and Korean ('KK') also participated in control experiments in those languages. However, since the native control participants in the QP-QP

[137] Within the scope of the present project, it was not possible to control this variable. Most L1 English-speaking learners of Japanese in the UK begin learning Japanese at university after having studied European languages from age 11. Most L1 Chinese- and Korean-speaking learners of Japanese study English as a foreign language prior to beginning Japanese. Thus, it was not feasible, for this study, to find adult learners of Japanese for whom Japanese was the sole foreign language they had studied.

[138] The quantification literature focuses on discussion of Mandarin Chinese, thus, ideally, all CJ participants (and native-speaking Chinese control participants) should have had Mandarin Chinese as their native dialect. However, in order to keep the number of participants in the Chinese-speaking groups as high as possible, it was decided to retain the data of participants whose native dialect was not Mandarin, or who were bilingual in two Chinese dialects. This was considered justifiable on the grounds that there is at least some evidence that judgements about quantifier scope are the same in non-Mandarin dialects of Chinese as in Mandarin Chinese. For example, informal investigation of the Phase 2A test sentences with (non-linguist) native speakers of Cantonese revealed that they tended to reject object-wide scope. In addition, discussion with Cantonese and Taiwanese linguists (Yuet Wah Lam, personal communication, July 2002; Nonie Chang, personal communication, June 2004) indicates that *Wh*-QP scope interpretations in those two dialects are as in Mandarin.

experiments were not always the same as those in the *Wh*-QP experiments, details of the native control groups are given with the details of each experiment.

None of the participants in the main study had taken part in any of the pilot-testing for Phase 1 or Phase 2. Appendix 3 lists all the participants with their age, gender, cloze test results (for the native Japanese and L2er participants) and Chinese dialect (for the native Chinese-speaking participants).

5.1.3. Proficiency task

The discussion of the pilot study (Phase 2A) in the previous chapter highlighted the need for a measure of the L2ers' proficiency in Japanese. The method chosen was a random cloze test. Every 7th word was removed from a passage taken from *Nihongo Journal* (2002).[139] A total of 42 blanks was inserted in place of the missing words. Participants were required to fill in one word per blank. The scripts were marked on an exact-word basis: only the exact word from the original text was accepted as a correct answer.[140] The cloze test is presented in Appendix 4.

This method of measuring L2 proficiency has been found to be reliable by language testing researchers (Jonz 1990; Oller 1979) and has been used in other L2 acquisition studies (e.g., Chen 1996; Montrul 1997; Slabakova 2001). The method has also been found to be reliable specifically for L2 Japanese by Shin (1983, 1987, cited in Ishida 1992: 160). However, some problems arise when using cloze tests for Japanese. First, as observed by Ishida (1992: 160), it is not obvious what should be considered a 'word', since spaces are not used between 'words' in written Japanese, unlike in languages using the Roman alphabet. Two potential problems in this regard are compound nouns, such as 電気製品 *denki-seihin* 'electrical goods', and verbs comprising a main verb and an aspectual verb, such as 食べている *tabete-iru* 'be eating (*lit*. eating-be)'. It was decided that elements such as these would be treated as two words in the present cloze test. This was explained to the participants, and illustrated with examples, before beginning the test.

A further problem arises because of the L1s of the learners in the present study. Written Chinese uses many of the same ideographs as Japanese, so L1 Chinese-speaking learners of Japanese can use their knowledge of Chinese to guess

[139] *Nihongo Journal* is a magazine for learners of Japanese.
[140] Answers were accepted as correct whether written in *kanji* (ideographs) or *hiragana* (phonetic script).

at the meanings of ideographs, while English and Korean learners have no such means of guessing unknown ideographs.[141] To attempt to neutralise this possible advantage for L1 Chinese-speaking learners of Japanese, *furigana* (phonetic superscript symbols) were added to all ideographs in the cloze test for which the Chinese meaning is the same as the Japanese meaning. It was thought that this would enable the English and Korean learners to at least 'sound out' the ideograph and potentially recognize its meaning from its phonetic value if not from its ideographic form. Thus, *furigana* were added to, for example, the two-ideograph compound 友人 because this means 'friend' in both Japanese and Chinese. The ideograph 探, on the other hand, was not glossed with *furigana*, because this ideograph occurs in Chinese only in two-ideograph compounds with meanings related to the verbs *visit* or *investigate*, while in Japanese, the single ideograph occurs in conjunction with verbal morphology indicated by phonetic Japanese syllables (e.g., 探す *sagasu*), with the meaning of 'to search for'. Since there is only an indirect correspondence between the meaning of 探 in Chinese and its use in Japanese, a learner who does not know the meaning in Japanese cannot easily guess it based on L1 knowledge, whether that L1 is Chinese, Korean or English.[142] While this method does not completely remove the reading advantage of L1 Chinese-speaking learners of Japanese, it was intended to at least be a step towards leveling the playing field.

The cloze test was completed by the L2 Japanese participants described in the previous section, as well as by 30 native speakers of Japanese (JJ).[143] The cloze test was presented simply as one of the four tasks making up the full task set; it was not presented specifically as a proficiency test. A maximum of 30 minutes was allowed for completion of the task (including five minutes' explanation time). In practice, the JJ participants completed the task within 10–15 minutes, and all learners finished

[141] Ideographs have been used in (South) Korean in the past, but at present they are not used in most written media. Thus, (South) Koreans currently of university age (the age of the participants in this study) do not generally have a broad knowledge of ideographs that would help them to guess meanings in a Japanese cloze test.

[142] One might ask why not just add *furigana* to all of the ideographs in the test. However, a learner's ability to read and understand Japanese ideographs is one measure of his or her proficiency. If all the ideographs were glossed with *furigana*, one element of the proficiency task would be lost.

[143] The JJ participants were university students and final-year high school students, all resident in Japan at the time of testing. The age range was 18 to 31, with a mean age of 20.

well within the 30 minute limit. The scores for each group are presented in Table 23.[144]

Table 23: Group scores on the cloze test (out of 42)

	lowest score	highest score	group mean
JJ (n=30)	12	33	22.4
EJ (n=29)	3	18	9.3
CJ (n=17)	3	20	11.5
KJ (n=38)	1	29	11.2

Since 12 out of 42 was the lowest score by a native speaker of Japanese, this score was taken as the minimum for a learner to be classed as 'advanced'. In other words, the 'intermediate' learners in Table 22 in the previous section scored between 1 and 11 on the cloze test, while the 'advanced' learners scored at least 12. Comparisons, by means of independent samples *t*-tests, of the mean proficiency task scores for each intermediate group with the advanced group of the same L1 confirm that the intermediate and advanced groups are significantly different: EJ int (n=20) *v.* EJ adv (n=9), $t(27)=-8.53, p<.001$; CJ int (n=7) *v.* CJ adv (n=10), $t(15)=-5.16, p<.001$; KJ int (n=23) *v.* KJ adv (n=15), $t(19.53)=-8.06, p<.001$.[145] See Appendix 4B for full statistical details.

5.2. Main study 1: QP-QP interpretation[146]

The main QP-QP study uses a modified version of the acceptability judgement task developed in Phase 2A (Chapter 4) to investigate quantifier interpretation in doubly-quantified Japanese sentences such as (186), by intermediate and advanced L2ers with three different L1s: English, Chinese and Korean.

186. Dareka-ga dono hon-mo yonda.
someone-NOM every book-QPt read
'Someone read every book.'

[144] Although some of the learner scores are very low, no learners were excluded from the study on the basis of their cloze test scores. Instead, distractors in the main studies were used to identify any learners whose proficiency may have been inadequate to the task. (See Sections 5.2 and 5.3.)
[145] 'Equal variances not assumed' values are used for the KJ comparison due to significant Levene's test result: $p<.05$.
[146] Some of the results of this main study are reported in Marsden (2004).

As already detailed (Chapter 2), a Japanese sentence such as (186) is unambiguous, allowing only a subject-wide scope interpretation 'there is some person such that that person read every book'. The logic of the present study rests on the fact that Chinese and (canonically-ordered SOV) Korean doubly-quantified sentences are similarly unambiguous (e.g., Aoun & Li 1993; Kim 1989), whereas English additionally allows an object-wide interpretation (i.e., in (186), 'for every book, there is some person who read it'). Thus, the L2 acquisition of the scope interpretation possibilities of Japanese sentences like (186) represents a different problem for learners whose L1 is Korean or Chinese compared with learners whose L1 is English. For L1 Korean- and L1 Chinese-speaking learners of Japanese, under Full Transfer/Full Access, the initial interlanguage grammar is target-like (i.e., disallows object-wide scope) due to L1 transfer. However, for L1 English-speaking learners of Japanese, the initial interlanguage grammar allows object-wide scope, as in English. Thus it is predicted that L1 English-speaking learners of Japanese will initially allow object-wide scope interpretations of Japanese sentences like (186), whereas L1 Korean- and L1 Chinese-speaking learners of Japanese will disallow object-wide scope from the outset.

The second prediction, based on Full Transfer/Full Access, is that advanced English-speaking learners of Japanese will be able to acquire target-like knowledge of the lack of object-wide scope in Japanese sentences like (186), despite the fact that acquisition of the relevant facts is a poverty-of-the-stimulus problem for L1 English-speaking learners of Japanese. Poverty of the stimulus arises because the target language input logically cannot contain any direct evidence to motivate induction of the lack of object-wide scope: although learners may hear sentences like (186) in subject-wide contexts, the fact that they do not encounter such a sentence in an object-wide context does not rule out the possibility of object-wide interpretations, given that the L1 grammar allows object-wide scope. Full Transfer/Full Access predicts that advanced learners may nonetheless be able to acquire native-like knowledge of the lack of object-wide scope due to subtle indirect evidence in the input triggering the re-setting of UG parameters to create a Japanese-like grammar (see Chapter 2).

Two different types of universal QP—*dono N-mo* 'every N' and *subete-no N* 'all the N'—are used in the investigation of these two predictions. As already discussed, object-wide scope is possible in English when the object quantifier is *every* (e.g., in *Someone read every book*), but it tends to be rejected when the object quantifier is *all* (*Someone read all the books*). In Japanese, Korean, and Chinese, object-wide scope is ruled out in canonical [S…O…] doubly-quantified sentences whatever universal quantifier modifies the object. The experiment aims to find out whether this difference in the scopal properties of specific quantifiers in native English transfers to the L1 English-L2 Japanese interlanguage.

To summarise, L1 Korean- and L1 Chinese-speaking learners of Japanese are predicted to demonstrate target-like interpretation of Japanese doubly-quantified sentences whether the object QP is *dono N-mo* 'every N' or *subete-no N* 'all the N', whatever the learners' level of Japanese proficiency. L1 English-speaking learners of Japanese, on the other hand, are predicted to initially allow (non-target-like) object-wide scope in Japanese sentences like (186), with *dono N-mo* 'every N' as object QP. However, when the object QP is *subete-no N* 'all the N', lower proficiency learners are predicted to reject object-wide scope, due to transfer of the properties of *all* to the interlanguage. Higher proficiency L1 English-speaking learners of Japanese are predicted to reject object-wide scope whatever the object QP. These predictions are expressed as hypotheses:

187. *QP-QP Hypothesis 1*: (L1 = Korean or Chinese)
Transfer: Lower (and higher) proficiency Korean-speaking and Chinese-speaking learners will reject non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences, regardless of the lexical content of the universally quantified object.

188. a. *QP-QP Hypothesis 2a* (L1 = English)
Transfer: Lower proficiency English-speaking learners will allow non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences when the universally quantified object is *dono N-mo* 'every N' but not when it is *subete no-N* 'all the N'.

b. *QP-QP Hypothesis 2b* (L1 = English)
Access: Higher proficiency English-speaking learners will reject non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences, regardless of the lexical content of the universally quantified object.

This section is organised as follows. Section 5.2.1 outlines the test design, Section 5.2.2, the procedure. The results are presented in Section 5.2.3, and discussion of the results with respect to the hypotheses, in Section 5.2.4.

5.2.1. Test design

Ten groups participated in the experimentation: the six L2 Japanese groups detailed in Section 5.1.2 (see Table 22); and four native control groups, described in Table 24.

Table 24: Native control participants in the QP-QP tasks

group	no.	age		details
		range	mean	
Japanese (JJ)	21	21–57	23	University students in Japan.
English (EE)	24	18–24	18	University students in UK.
Chinese (CC)	24	20–36	25	University students resident in UK.
Korean (KK)	26	21–27	22	University students in Korea.

The native control participants had all had exposure to languages other than their native language through pre-university and (in most cases) university education in their home countries. However, none of the Japanese, English or Korean participants were bilingual from early childhood, and none had lived outside their home countries. The mean length of residence in the UK for the native Chinese group was 15 months (range: 2 months–12 years). Twenty-two of the 24 native Chinese participants gave Mandarin or a variant of Mandarin as their native dialect or their dominant dialect if bilingual in more than one Chinese dialect. One Chinese participant was a native Cantonese speaker, and one a native Taiwanese speaker.[147] (See Section 5.1.2 for comments on the treatment of Chinese dialects.)

The test instrument was an adapted version of the picture-based acceptability judgement task used in Phase 2A. Four repeated measures variables (189–192) were manipulated in the task. In addition to the group variable, the variables of object QP (189) and scope (190) are of key importance to the hypotheses given above (see (187–188)):

189. *Variable 1: object QP*
 dono N-mo 'every N' v. *subete-no N* 'all the N'

190. *Variable 2: scope*
 subject-wide scope v. object-wide scope

In addition, the variables of subject QP and scrambling were also manipulated:

[147] Neither produced response patterns that stood out from the native Mandarin-speaking participants' response patterns.

191. *Variable 3: subject QP*
dareka 'someone' v. a numerically quantified noun (NumP)

192. *Variable 4: scrambling*
canonical word order (SOV) v. scrambled word order (OSV)

The subject QP variable was investigated in Phase 2A and found to produce no significant main effect or interaction effect. Thus, an option for simplifying the test design would have been to keep the subject QP constant across all test sentences. However, since it had not been possible to pilot the test with native Chinese- and Korean-speaking learners of Japanese or with native Korean or native Chinese speakers, it was decided to retain both subject QP types just in case one or other of them proved problematic in the unpiloted groups. The scrambling variable was reintroduced (having been investigated in Phase 1) in order to discover whether learners of Japanese differentiate between canonical and scrambled Japanese sentences in the same way as native Japanese speakers do. The expected native Japanese judgements for subject-wide (S>O) and object-wide (O>S) scope on canonical (SOV) and scrambled (OSV) sentences are schematised in (193).

193. a. word order = SOV, scope = S>O: accept
b. word order = SOV, scope = O>S: reject
c. word order = OSV, scope = S>O: accept
d. word order = OSV, scope = O>S: accept

Supposing that a learner rejects object-wide scope in canonical sentences (193b), as in native Japanese, but also rejects object-wide scope in scrambled sentences (193d). This could indicate a tendency to always reject object-wide scope. Alternatively, a learner might correctly reject object-wide scope in canonical sentences (193b), but also reject subject-wide scope in scrambled sentences (193c). This could indicate a tendency to always reject inverse scope. Thus, if learners who correctly reject object-wide scope on canonical SOV QP-QP sentences also accept both subject-wide and object-wide scope on scrambled QP-QP sentences, this would provide stronger evidence of a native-like interlanguage grammar than rejection of object-wide scope in canonical sentences alone.

Twelve test types were created, as shown and exemplified in Table 25.

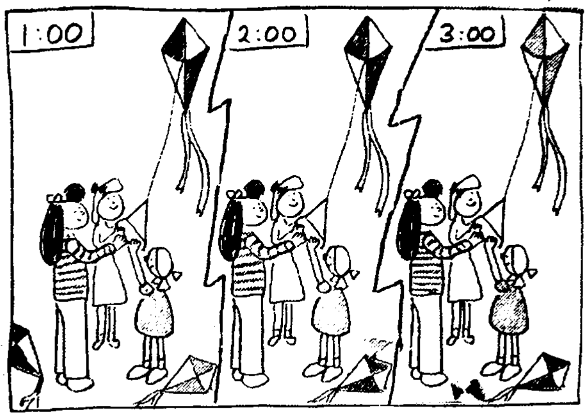
Table 25: Test types, QP-QP experiment

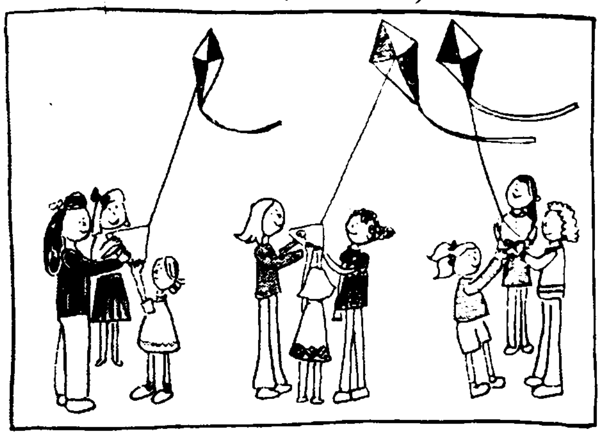
type	variables				example
	subject QP	object QP	word order	scope	
1a	*dareka* 'someone'	*dono-N mo* 'every N'	SOV	S>O	Dareka-ga dono neko-mo nadeta. someone-NOM every cat-QPt stroked 'Someone stroked every cat.'
1b				O>S	
1c			OSV	S>O	Dono neko-mo dareka-ga nadeta. every cat-QPt someone-NOM stroked 'Someone stroked every cat. (*scrambled*)'
1d				O>S	
2a		*subete-no N* 'all the N'	SOV	S>O	Dareka-ga subete-no suutukeesu-o hakonda. someone-NOM all-GEN suitcase-ACC carried 'Someone carried all the suitcases.'
2b				O>S	
3a	NumP	*dono-N mo* 'every N'	SOV	S>O	Sannin-no onnanoko-ga dono tako-mo ageta. three-GEN girl-NOM every kite-QPt flew 'Three girls flew every kite.'
3b				O>S	
3c			OSV	S>O	Dono tako-mo sannin-no onnanoko-ga ageta. every kite-QPt three-GEN girl-NOM flew 'Three girls flew every kite. (*scrambled*)'
3d				O>S	
4a		*subete-no N* 'all the N'	SOV	S>O	Hutari-no onnanoko-ga subete-no mado-o aratta. two-GEN girl-NOM all-GEN window-ACC washed 'Two girls washed all the windows.'
4b				O>S	

Types 1a–b, 2a–b, 3a–b, and 4a–b are the same as the types used in Phase 2A. Types 1c–d and 3c–d are scrambled variants of 1a–b and 3a–b. Scrambled variants of 2a–b and 4a–b (the types with *subete-no N* 'all the N' as object QP) were not created due to concern that the test would be too long.

As in the pilot-testing, the scope variable was manipulated by means of a picture context for each sentence. To exemplify, the subject-wide and object-wide pictures for the Types 3a–d sentences in Table 25 are shown below (the kites in each picture were differentiated by colour):

194. a. *S>O scope context* (3a & 3c): b. *O>S scope context* (3b & 3d):

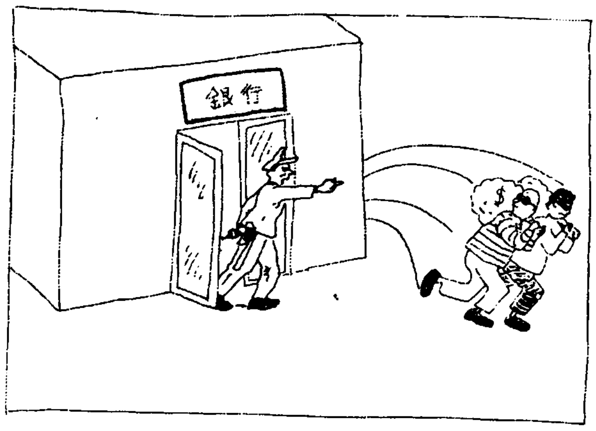




Five test tokens were created for each type. Using five tokens instead of four (as in Phase 2A), the consistency of an individual participant's responses on a particular type can be measured slightly more strictly. For example, the criterion for 'consistent rejection' of a test type can be set at 80% (at least four tokens out of five with a 'rejection' rating) instead of the 75% (at least three tokens out of four with a 'rejection' rating) used in Phase 2A. This was intended to increase the validity of analyses of individual consistency.

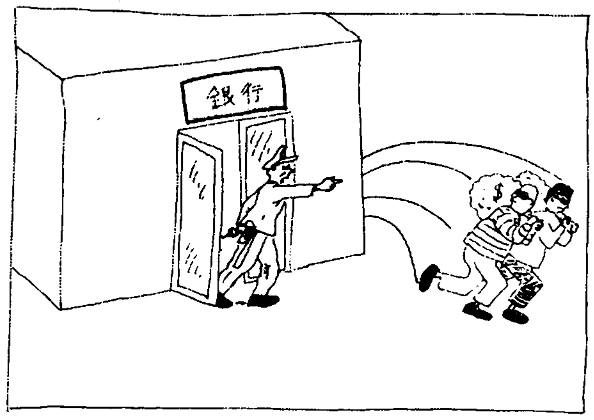
The test battery was divided into two sets. Set 1 ('QP-QP Task 1') contained the sentences for Types 1a–b, 3c–d and 4a–b. Set 2 ('QP-QP Task 2') contained the sentences for Types 1c–d, 2a–b and 3a–b. Thus each task included 30 test items. Fourteen distractors were added to each test set. It was considered that 14 distractors were enough, because the different test types within each task also provide a certain amount of distraction from each other. As in Phase 2A, the distractor items were designed to blend in with the test items. Their structure was either [SOV.PAST] or [OSV.PAST]. Every distractor item contained at least one quantified NP. The distractor items used the same pictures in both QP-QP Task 1 and QP-QP Task 2. Four of them also had the same sentence in the two tasks. The remaining 10 had different sentences in each task. The difference lay in the word order: if a canonical word order was used in QP-QP Task 1, the scrambled word order was used in QP-QP Task 2. An example is given in (195), where (195a) shows the canonical word order and (195b) the scrambled word order.

195. a. *QP-QP Task 1, Distractor D01*:



Keisatukan-ga hutari-no dorobou-o mita
police officer-NOM two.GEN thief-ACC saw
'A policeman saw two thieves.'

b. *QP-QP Task 2, Distractor D01*:



Hutari-no dorobou-o keisatukan-ga mita
two.GEN thief-ACC police officer-NOM saw
'A policeman saw two thieves. (*scrambled*)'

Five of the 10 distractors that differed between the two tasks had a canonical (SOV) word order and five had a scrambled (OSV) word order in each task. All 14 distractor items in each task were intended to be unambiguous: in seven of them the picture matched the sentence, and in the other seven, the picture and sentence did not match. The full Japanese test battery is presented in Appendix 5A.

The first and last item of each task was a distractor. The remaining 12 distractors and the 30 test items were mixed randomly to form 'Order 1' of each task. 'Order 2' was the reverse of Order 1. Some participants completed Order 1 for each task, others Order 2. This was intended to minimise any potential effect of the order in which the test items appeared.[148]

[148] Conditions at the institutions where the tasks were conducted sometimes dictated that all participants take the tests at one time, thus only one test order could be used. This happened for the native Korean and Chinese groups, and for the native Japanese group on QP-QP Task 1, although both orders were used with the native Japanese group on QP-QP Task 2.

English, Chinese and Korean test batteries were also created, in order to collect native control data. The English and Chinese test batteries did not include the scrambled sentence types (Types 1c, 1d, 3c and 3d) because these two languages do not have scrambling. The test design for these two languages is presented and exemplified in Table 26. (The English translations of the Chinese illustrate the design used for the English version of the task.)

Table 26: Test types for Chinese and English versions of QP-QP experiment

type	variable			example
	subject QP	object QP	scope	
1a	*mouren* 'someone'	*mei-CL N* 'every N'	S>O	Mouren fumole mei-zhi mao. someone stroked every-CL cat 'Someone stroked every cat.'
1b			O>S	
2a		*suoyoude N* 'all the N'	S>O	Mouren xiedaile suoyoude xinglixiang. someone carried all suitcase 'Someone carried all the suitcases.'
2b			O>S	
3a	NumP	*mei-CL N* 'every N'	S>O	San-ge nuhai fangle mei-zhi fengzheng. three-CL girl flew every-CL kite 'Three girls flew every kite.'
3b			O>S	
4a		*suoyoude N* 'all the N'	S>O	Liang-ge nuhai gingxile suoyoude chuanghu. two-CL girl washed all window 'Two girls washed all the windows.'
4b			O>S	

As is evident from Table 26, the Chinese QP *mei-CL N* is used in the present task as the equivalent of English *every N*, and Chinese *suoyoude N* is used as the equivalent of *all the N*. This follows Lee, Yip & Wang (1999a, 1999b). However, as noted in Chapter 2 with respect to Japanese *dono N-mo* 'every N' and *subete-no N* 'all the N', the English translations of these QPs are approximations that are not intended to imply exact semantic correspondence. QP-QP Task 1 in Chinese and English contained Types 1a–b and 4a–b from Table 26, as well as translations of the 14 distractor items used in the Japanese version of QP-QP Task 1. QP-QP Task 2 in Chinese and English contained Types 2a–b and 3a–b from Table 26 as well as translations of the 14 distractor items used in the Japanese version of QP-QP Task 2. The distractors with a scrambled word order in Japanese were usually passivized in the English and Chinese tasks. The full Chinese and English test battery is presented in Appendix 5B.

The Korean QP-QP task design also differed from the Japanese design because only one type of object QP was investigated: *motun N*, which is syntactically and semantically closer to Japanese *subete-no N* 'all the N' than to Japanese *dono N-mo* 'every N'.[149] For example, while *dono N-mo* 'every N' cannot c-command a *wh*-phrase, as shown in (196), *subete-no N* and *motun N* produce no such ungrammaticality effect, as illustrated in (197) and (198) (examples adapted from Tomioka 2004).

196. ??Dono gakusei-mo nani-o yonda no?
every student-QPt what-ACC read Q
'What did every student read?'

197. Subete-no gakusei-ga nani-o yonda no?
all-GEN student-NOM what-ACC read Q
'What did all the students read?'

198. Motun haksayng-i mues-ul ilkess ni?
all student-NOM what-ACC read Q
'What did all the students read?'

In addition, *motun N* tends to have a collective interpretation rather than a distributive interpretation (Kook-Hee Gill, personal communication, March 2004), just like Japanese *subete-no N* (see Chapter 2) and English *all the N*. Therefore, *motun N* is translated as 'all the N', although, as with other translations of quantifiers, this is not intended to indicate exact semantic correspondence of Korean *motun N* with English *all the N*.

The Korean test design is given in Table 27. The Korean test types are labeled 'K1a', 'K1b', etc., instead of '1a', '1b', etc., because they do not correspond exactly to any of the test sentences in the Japanese, Chinese and English versions of the test. For example, the design of Type K1a corresponds to the design of the Japanese (and Chinese and English) Type 2a (i.e., subject QP = 'someone', object QP = 'all the N', scope = S>O), but the vocabulary used in the Korean Type K1a

[149] Korean also has the universal QP *enu N-to* 'every N', which is morphologically, syntactically and semantically equivalent to Japanese *dono N-mo*. *Enu N-to* was not included in the Korean version of the experimentation due to insufficient knowledge of this QP at the time of having the opportunity to work with native Korean test participants. To gain at least pre-experimental data on *enu N-to*, Korean versions of Types 1a–b and 3a–b were created (with *enu N-to* as the object QP) and presented to two native Korean linguists. They reported that object-wide scope was unacceptable, as expected following the theoretical literature. (Beck & Kim (1997: 374) discuss one example with *enu N-to* 'every N' as object QP; other sources, such as Kim 1989, discuss only *nwukwuna* 'everyone'.)

tokens was the same as the vocabulary used in Japanese Type 1a. Consequently, the results for the native Korean QP-QP task are not compared directly with the other native results in Sections 5.2.3.2–3; they are presented separately in Section 5.2.3.4.

Table 27: Test types for Korean version of QP-QP experiment

type	variable			example
	subject QP	word order	scope	
K1a	*nwukwunka* 'someone'	SOV	S>O	Nwukwunka-ka motun koyangi-lul ssutatumessta. someone-NOM all cat-ACC stroked 'Someone stroked all the cats.'
K1b			O>S	
K1c		OSV	S>O	Motun koyangi-lul nwukwunka-ka ssutatumessta. all cat-ACC someone-NOM stroked 'Someone stroked all the cats. (scrambled)'
K1d			O>S	
K2a	NumP	SOV	S>O	Sey-myeng-uy sonya-ka motun yen-ul nallessta. three-CL-GEN girl-NOM all kite-ACC flew 'Three girls flew all the kites.'
K2b			O>S	
K2c		OSV	S>O	Motun yen-ul sey-myeng-uy sonya-ka nallessta. all kite-ACC three-CL-GEN girl-NOM flew 'Three girls flew all the kites. (*scrambled*)'
K2d			O>S	

QP-QP Task 1 in Korean contained Types K1a–b, K2c–d from Table 27, as well as translations of the 14 distractor items used in the Japanese version of QP-QP Task 1. QP-QP Task 2 in Korean contained Types K1c–d, K2a–b from Table 27 as well as translations of the 14 distractor items used in the Japanese version of QP-QP Task 2. The full Korean test battery is presented in Appendix 5C.

5.2.2. Procedure

The pictures and their corresponding sentences were projected onto an overhead projector screen. Each picture was presented without the sentence for 10 seconds, then the sentence was shown in written form underneath the picture for a further 15 seconds. At the same time as the sentence was revealed, it was also presented aurally, using an audio tape recorded by a native speaker of the language of the test. Participants were asked to judge how well each picture matched the corresponding sentence by selecting one of four numerical ratings, –2, –1, +1, +2, or a 'can't decide' option. It was explained that the negative ratings '–2' and '–1' indicated degrees of rejection of the picture-sentence match, and the positive ratings indicated degrees of acceptance.[150] The scale on the answer sheet is illustrated in (199). (The complete answer sheet is given in Appendix 5D in English and Appendix 5E in Japanese.)

199. *Answer sheet layout for QP-QP Tasks 1 and 2*:

(Does the sentence match the picture?)

	No, definitely not			Yes, perfectly		Can't decide
1	−2	−1	+1	+2		X
2	−2	−1	+1	+2		X

The four-point scale contrasts with the five-point scale (−2, −1, 0, +1, +2) used in Phase 2A. As noted in the previous chapter, the '0' response on the five-point scale is ambiguous: it is unclear whether participants select '0' because they believe the picture-sentence match is as acceptable as it is unacceptable, or because they do not know how to judge the picture-sentence match. With the four-point scale used in the present study, the participants did not have the option of judging a picture-sentence match equally acceptable and unacceptable. However, it was decided that omission of this option would not be problematic because the rate of selection of '0' in Phase 2A had been very low (3.19% of all responses), and at least some of these were likely to have been intended to convey a 'can't decide' response. The four-point scale with a 'can't decide' option was considered, overall, to be clearer for participants than the five-point scale.[151]

Instructions on how to do the test were given on the test sheet and were also explained orally. The instructions included six example test items, shown in (200).[152]

---

[150] It was determined in the development stages (Chapter 4) that the use of a negative-positive scale was more appropriate to judgements about scope interpretation, which may sometimes be vague, than a scale allowing only a binary choice between 'acceptable' and 'not acceptable'.

[151] The four-point scale could have been expressed as a scale of 0, 1, 2, 3. This would have the advantage of an equal difference of one point between each point on the scale, whereas the scale used has a difference of two points between −1 and +1 and one point between −2 and −1, and +1 and +2. The 0–3 scale was not adopted, however, because of comments from participants in Phase 2A about the user-friendliness of the association of negative numbers on the rating scale with picture-sentence mismatches and positive numbers with picture-sentence matches. For the analysis, the scale −2, −1, +1, +2 is transformed to 0, 1, 2, 3 (see Section 5.2.3).

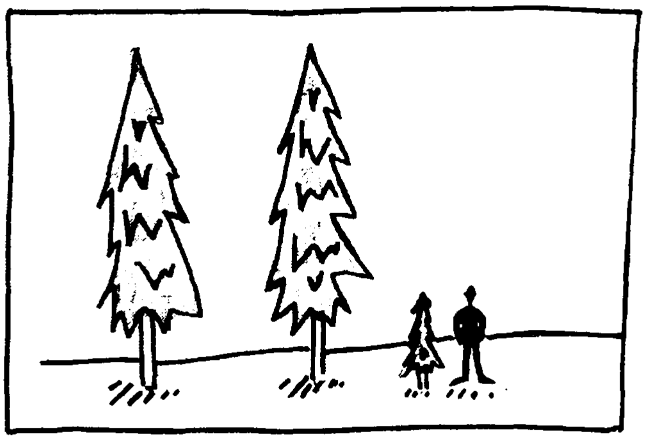
[152] For convenience, an English translation of the Japanese text in the picture for Example 1 (200a) is included underneath the picture. In the actual demonstration of the examples, there was no translation. English, Chinese and Korean versions of the examples in (200) were used for the English, Chinese and Korean QP-QP tasks.

200. a. *Example 1:*

Translation: 'Ha, ha, ha...'

Dareka-ga waratta.
Someone-NOM laughed.
'Someone laughed.'

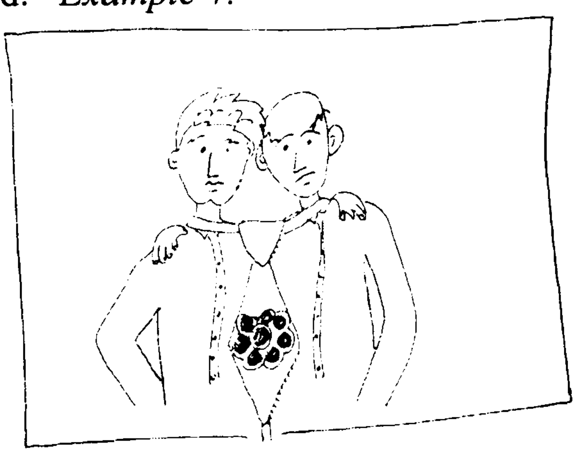
b. *Example 2:*



Dono ki-mo taka-katta
Every tree-QPt tall-PAST.
'Every tree was tall.'

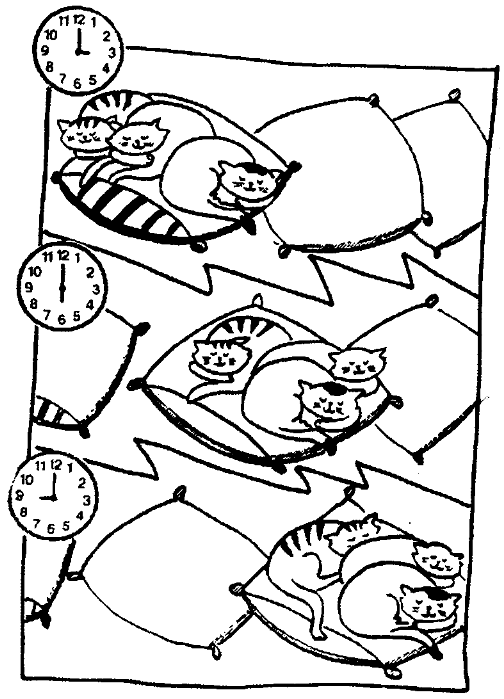
c. *Example 3*:

d. *Example 4*:

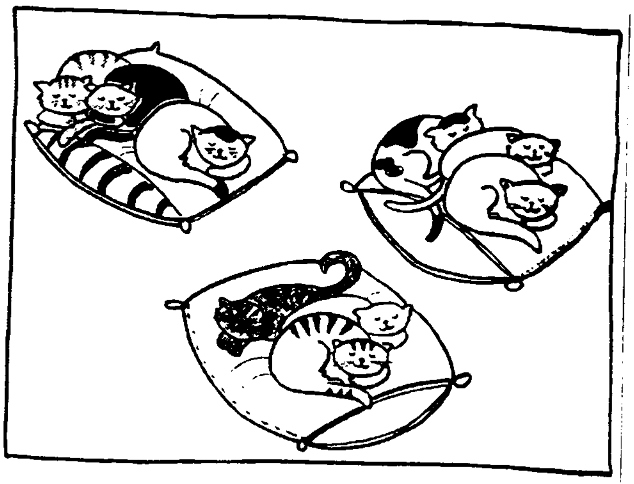


Sentence for examples 3–4:
Hutari-no otoko-ga onazi nekutai-o sita.
two-GEN man-NOM same necktie-ACC wore.
'The two men wore the same necktie.'

e. *Example 5*:



f. *Example 6*:



Sentence for examples 5–6:
Onazi san-biki-no neko-ga sorezore-no kussyon-no ue ni neta.
same three-CL-GEN cat-NOM each-GEN cushion-GEN top on slept
'The same three cats slept on each cushion.'

As in the pilot-test stages, the examples were presented one by one with time for participants to select a rating, then ask questions or discuss the rating between each example. The broad aim was to introduce participants to the rating scale. In addition, the examples were designed to prepare the participants for the types of sentences to be encountered in the actual task. Thus, all the example sentences are in the past tense, just like the real test items. As in Phase 2A, participants were instructed to think of each picture-sentence pair as referring to events that occurred 'yesterday'. The specific aim of each example was as follows.[153] Example (1) in (200a) was designed to introduce participants to the idea that the existential *dareka* 'someone' might indicate a person illustrated in the picture. The purpose of Example 2 was to exemplify a sentence using *dono N-mo* 'every N', and to point out that participants should always think of 'every' within the context of the picture, not within the context of the whole world. Examples 3 and 4 (200c–d) were intended to prepare participants for encountering the same sentence with different picture contexts. Example 4 shows how the interpretation of a picture can be unexpected but nonetheless possible. Examples 5 and 6 (200e–f) were intended to illustrate the

[153] The examples are very similar to those used in Phase 2A. Only Example 1 is brand new. In Example 2 the adjective is *taka-* 'tall' instead of *ooki-* 'big', because some Phase 2A participants had wondered whether the trees could really be described as 'big' due to their slenderness. The wording of Examples 5 and 6 is also slightly amended.

importance of looking carefully at the details of the pictures and noticing, as in Example 5, that the three cats are the same in each section of the picture; or, as in Example 6, that nine different cats are illustrated. As in the pilot tests, it was stressed that there was no 'right' or 'wrong' rating, and participants should respond according to their intuitions. Finally, it was pointed out that all the sentences were grammatically correct, so the task was not about identifying grammar errors.

5.2.3. Results

This section presents details of the distractor results first (Section 5.2.3.1), since these were used to determine whether any participants' data should be excluded from the analysis. Section 5.2.3.2 presents the results for the non-scrambled test sentences in Japanese, English, Chinese and L2 Japanese; Section 5.2.3.3 presents the results for the scrambled test sentences in Japanese and L2 Japanese; and Section 5.2.3.4 presents the results for the Korean version of the task.

The data were coded using Excel. For the analysis, the rating scale of '−2, −1, +1, +2' was transformed to '0, 1, 2, 3' (see footnote 151). The transformed scale is used throughout this section, with numbers from the original scale given in brackets, where relevant. For example, 'a rating of 0 (−2)' means, 'a rating coded as 0 for the analysis but which was −2 on the actual rating scale'. Responses of 'can't decide' were treated as missing answers and excluded from the main analysis.[154] However, in

[154] The number of 'can't decide' responses and the number of missing (or illegible) responses were low in the data retained for analysis in Sections 5.2.3.2–4, as shown in Table I.

Table I: Number of 'can't decide' and missing responses by participant group

	no. of responses		total
	'can't decide'	missing	possible[a]
JJ (n=20)	4	1	1200
EE (n=24)	5	0	960
CC (n=20)	1	1	800
KK (n=22)	2	0	880

	no. of responses		total
	'can't decide'	missing	possible[a]
EJ int (n=18)	8	1	1080
EJ adv (n=9)	0	0	540
CJ int (n=6)	0	3	360
CJ adv (n=9)	0	1	540
KJ int (n=20)	6	4	1200
KJ adv (n=15)	0	0	900

[a] 'Total possible' = no. participants x no. test items on QP-QP Tasks 1 and 2 (60 for Japanese version of tasks; 40 for English, Chinese and Korean versions of tasks)

One individual in the intermediate EJ group selected 'can't decide' four times (one Type 1c token, one Type 3c token, and two Type 4a tokens). The other instances of 'can't decide' were due to individuals selecting this response option once or twice. There were no more than two missing answers in any individual's response set.

the analysis of the distractor items, 'can't decide' responses were counted as incorrect answers.

The ratings for the five tokens for each test type were averaged for each participant. On the transformed scale, mean ratings of below 1.5 are considered to indicate rejection of that test type, and mean ratings of above 1.5, acceptance (since, on the original scale, –2 and –1 indicated rejection, and +2 and +1, acceptance). Group means were computed and analysed using the statistics package SPSS. All statistical details are presented in Appendix 5F.

5.2.3.1. Distractor items

The responses to the distractor items were examined first, in order to find out whether to exclude any participants' data from the analysis due to possible lack of understanding of the task. The native Japanese mean ratings and standard deviations for the distractor items are presented in Table 28 (QP-QP Task 1) and Table 29 (QP-QP Task 2). (Distractors D10, D11, D12 and D13 were the same in both tasks. The other distractor items were different in each task, as explained in Section 5.2.2.)

Table 28: Native Japanese ratings on QP-QP Task 1 distractor items (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10	D11	D12	D13	D14
Mean	3	1.6	2.52	2.86	2.71	0.55	0.14	2.86	0.1	0.05	0.1	2.14	2.57	0.57
SD	0	1.27	0.81	0.36	0.46	0.89	0.48	0.48	0.3	0.22	0.3	0.85	0.81	1.08

Table 29: Native Japanese ratings on QP-QP Task 2 distractor items (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10	D11	D12	D13	D14
Mean	2.95	0.81	2.33	2.86	2.57	0.35	0.24	2.95	0.14	0.05	0.1	2.43	2.62	0.19
SD	0.22	1.08	0.86	0.36	0.81	0.59	0.44	0.22	0.48	0.22	0.3	0.68	0.5	0.68

All the distractors were intended to be unambiguous, so mean ratings were expected to be close to 0 (–2), indicating rejection of the picture-sentence match, or close to 3 (+2), indicating acceptance of the picture-sentence match. Distractors with a mean rating of between 1 (–1) and 2 (+1) and a standard deviation greater than 1 were considered to be potentially unreliable. Only item D02 in QP-QP Task 1 (Table 28) falls into this category (mean = 1.6, SD = 1.27). In fact, this is the scrambled version of one of the unreliable distractors in Phase 2A (Distractor D06, discussed in in Chapter 4). The canonical version (i.e., the exact same item as used in Phase 2A) was used as item D02 in QP-QP Task 2, and Table 29 shows it had a mean rating of close to 1 (0.81) and a high standard deviation (1.08). The relevant sentence is given in (201) in the canonical word order.

201. Keisatukan-ga hutari-tomo dorobou-o mita.
police officer-NOM both thief-ACC saw
AMBIGUOUS: 'Both police officers saw the thieves.'
'The police officer(s) saw both thieves.'

This item was included in the main study by error. Since it had already been found to be unreliable in Phase 2A, due to genuine ambiguity, responses to item D02 in QP-QP Tasks 1 and 2 were ignored when considering participants' performance on the distractor items. All the remaining distractors in both tasks were considered reliable.

The criterion for inclusion of a participant's data in the analysis was set at a minimum of 23 correct distractor answers out of the total of 26 (across both tasks) with no more than two incorrect answers on either Task 1 or Task 2. 'Correct' is defined as a rating of 0 (–2) or 1 (–1) on distractors where the picture did not match the sentence (D06, D07, D09, D10, D11 and D14 in both tasks), and 2 (+1) or 3 (+2) on distractors where the picture matched the sentence (D01, D03, D04, D05, D08, D12 and D13 in both tasks). By this criterion, the data of one native Japanese participant were excluded from the analysis. Among the learner groups, incorrect distractor responses led to the exclusion of the data of two participants in the intermediate EJ group, one from the intermediate CJ group, one in the advanced CJ group, and two in the intermediate Korean group. A further intermediate KJ participant was excluded due to a spoiled answer sheet. The resulting group sizes for the analysis were: native Japanese, 20; intermediate EJ, 18; advanced EJ, 9; intermediate CJ, 6; advanced CJ, 9; intermediate KJ, 20; advanced KJ, 15.

Some unreliable distractors were also identified in the English, Chinese and Korean versions of the QP-QP tasks: one in the English version, four in the Chinese version and two in the Korean version.[155] These items were ignored when

[155] The unreliable items were as follows: in English, item D06 in Task 1 (mean = 1.21, SD = 1.02); in Chinese, item D04 in Task 1 (mean = 1.50, SD = 1.18) and Task 2 (mean = 1.88, SD = 1.12), and item D06 in Task 1 (mean = 1.21, SD = 1.14) and Task 2 (mean = 0.96, SD = 1.00); and in Korean, item D04 in Task 1 (mean = 1.50, SD = 0.99) and Task 2 (mean = 1.50, SD = 1.10). Comments on the reasons for the lack of consensus in judgements about these items appear with the relevant items in Appendices 5B (Chinese and English) and 5C (Korean).

considering whether any of the control participants' data should be excluded due to incorrect distractor responses. Thus there were 27 distractors remaining for consideration in the native English data, 24 in the native Chinese data, and 26 in the native Korean data. In English and Korean, the criterion for inclusion of a participant's data in the analysis was the same as in the Japanese task: a minimum of 23 correct distractor answers (out of 27 in English, 26 in Korean) with no more than two incorrect answers on either Task 1 or Task 2. In Chinese, the criterion for inclusion was set at a minimum of 22 correct distractor answers out of 24, with no more than one incorrect answer on either Task 1 or Task 2. On this basis, four native Chinese and four native Korean participants were excluded. There were no exclusions from the native English data. The resulting sizes of the control groups were: native English, 24; native Chinese, 20; native Korean, 22.

5.2.3.2. Results for non-scrambled QP-QP sentences in native Japanese, English and Chinese, and L2 Japanese

The mean ratings and standard deviations for Types 1a–b, 2a–b, 3a–b and 4a–b (the non-scrambled test types) are presented in Table 30.[156]

[156] Prior to calculation of the group means, the native Japanese mean scores were checked for each test item in order to find out whether any test item had an anomalously high or low value and was therefore potentially unreliable. On the subject-wide scope items (Types 1a, 2a, 3a and 4a), which were expected to be acceptable and thus yield mean scores close to 3 (+2), the lowest mean score was 1.80 on item 1.1a. All other mean scores on these four types were above 2.2. Thus, all the subject-wide scope items except item 1.1a were highly acceptable, as expected. Since the mean score of 1.80 for item 1.1a also indicates acceptance of this item, even though it is only slightly above the mid-point score of 1.5, the subject-wide scope items were considered to be unproblematic. The object-wide scope items (Types 1b, 2b, 3b and 4b) were also unproblematic: they were expected to be unacceptable and thus yield mean scores close to 0 (–2). In fact, the highest mean score was 1.00 on item 4.4b, thus all object-wide scope items clearly fell into the category of 'unacceptable'.

Table 30: Mean ratings and standard deviations for Types 1a–b, 2a–b, 3a–b and 4a–b (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

Group	Type 1a[a] (S>O) Mean (SD)	Type 1b[a] (O>S) Mean (SD)	Type 2a[b] (S>O) Mean (SD)	Type 2b[b] (O>S) Mean (SD)	Type 3a[c] (S>O) Mean (SD)	Type 3b[c] (O>S) Mean (SD)	Type 4a[d] (S>O) Mean (SD)	Type 4b[d] (O>S) Mean (SD)
JJ (n=20)	2.26 (0.70)	0.69 (0.72)	2.54 (0.54)	0.60 (0.54)	2.86 (0.27)	0.70 (0.62)	2.70 (0.31)	0.73 (0.66)
EJ int (n=18)	2.74 (0.28)	1.61 (0.63)	2.88 (0.16)	1.76 (0.67)	2.81 (0.27)	1.92 (0.77)	2.80 (0.29)	1.32 (0.71)
EJ adv (n=9)	2.82 (0.39)	1.38 (1.12)	2.98 (0.07)	1.16 (0.88)	2.93 (0.14)	1.31 (1.02)	2.96 (0.09)	1.04 (0.82)
CJ int (n=6)	2.07 (0.98)	1.19 (0.77)	2.07 (1.04)	0.60 (0.55)	2.17 (0.85)	0.90 (0.60)	2.00 (0.98)	0.63 (0.61)
CJ adv (n=9)	2.87 (0.17)	1.07 (0.78)	2.82 (0.37)	0.71 (0.44)	2.89 (0.20)	0.82 (0.53)	2.89 (0.20)	0.82 (0.60)
KJ int (n=20)	2.58 (0.45)	1.24 (0.71)	2.66 (0.41)	0.79 (0.70)	2.59 (0.49)	0.92 (0.58)	2.72 (0.46)	0.70 (0.66)
KJ adv (n=15)	2.60 (0.35)	0.64 (0.71)	2.92 (0.22)	0.51 (0.39)	2.85 (0.22)	0.75 (0.59)	2.88 (0.20)	0.41 (0.45)
EE (n=24)	2.68 (0.36)	1.96 (0.73)	2.98 (0.68)	0.93 (0.67)	2.91 (0.14)	1.74 (0.94)	2.95 (0.11)	0.85 (0.63)
CC (n=20)	2.26 (0.49)	0.27 (0.33)	2.60 (0.69)	0.41 (0.33)	2.36 (0.75)	0.45 (0.55)	2.22 (0.61)	0.43 (0.46)

[a] Types 1a–b: subject = *dareka/mouren/someone*; object = *dono N-mo/mei-CL N/every N*
[b] Types 2a–b: subject = *dareka/mouren/someone*; object = *subete-no N/suoyoude N/all the N*
[c] Types 3a–b: subject = NumP; object = *dono N-mo/mei-CL N/ every N*
[d] Types 4a–b: subject = NumP; object = *subete-no N/suoyoude N/all the N*

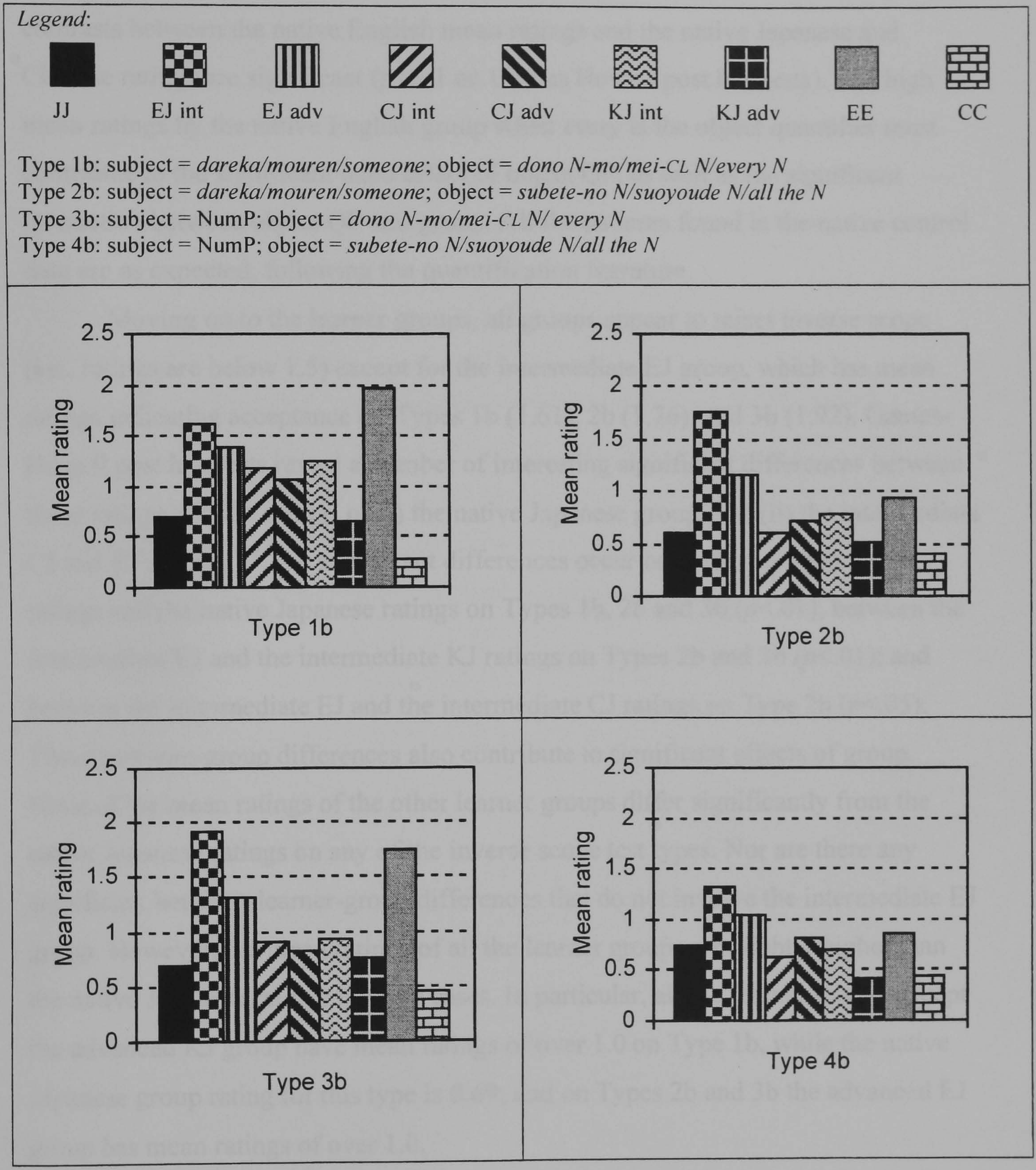
A repeated measures ANOVA (*subject QP x object QP x scope x group*) run on the data yields a number of significant effects.[157] The main effects of object QP ($F(1,132) = 23.10, p<.001$), scope ($F(1,132) = 827.15, p<.001$), and group ($F(8,132) = 15.61, p<.001$) are significant. The subject QP variable does not have a significant main effect ($F(1,132) = .01, p=.911$), although some of its interactions are significant with probabilities of just below .05: subject QP x object QP, $F(8,132) = 3.94, p<.05$; subject QP x scope, $F(8,132) = 4.39, p<.05$; subject QP x object QP x group, $F(8,132) = 2.37, p<.05$. Interactions of the other variables are highly significant: object QP x group, $F(8,132) = 7.11, p<.001$; object QP x scope, $F(8,132) = 50.77, p<.001$; scope x group, $F(8,132) = 4.55, p<.001$; object QP x scope x group, $F(8,132) = 7.72, p<.001$.

[157] See Appendix 5F for full details. Levene's test for homogeneity yielded significant results ($p>.05$) for Types 1a, 1b, 2a, 3a and 4a. Thus homogeneity of variance cannot be assumed across all the data. Consequently, *F*-ratios and *p*-values may not be as reliable as if homogeneity of variance were assured. Games-Howell tests are used for post hoc between-group analyses, since this procedure is more accurate when sample sizes and population variances are unequal (Field 2000: 276).

The sources of the key significant effects are generally clear from the data in Table 30. First, the significant effect of scope must arise from the fact that mean ratings on the linear scope test types (Types 1a, 2a, 3a and 4a) are all at least 2.00 or higher, whereas mean ratings on the inverse scope test types (Types 1b, 2b, 3b and 4b) range from 0.27 to 1.96. In other words, linear scope interpretations of the test sentences were acceptable to all groups; however, on the inverse scope test sentences, acceptability varied by group and by object QP.[158] This variation can be seen clearly in Figure 2, which illustrates the mean ratings on the inverse scope test types.

[158] Hereafter, discussion in this section refers only to the inverse (O>S) scope test items.

Figure 2: Mean ratings for object-wide scope on SOV QP-QP sentences
(<1.5 = 'unacceptable'; >1,5 = 'acceptable')



From the charts in Figure 2, it is clear that inverse scope interpretations are considerably more acceptable to some groups than others. This contributes to the significant main effect of group. Considering the native control groups first, the native English group stands out as having high mean ratings on Types 1b (1.96) and 3b (1.74), indicating that inverse scope is acceptable when the object QP is *every N*. However, on Types 2b and 4b (object QP = *all the N*), the native English ratings are

0.93 and 0.85, respectively, indicating that inverse scope is not acceptable. In native Japanese and native Chinese, inverse scope is not acceptable on any of the four types, with mean ratings no higher than 0.73 (JJ, Type 4b). On Types 1b and 3b, the contrasts between the native English mean ratings and the native Japanese and Chinese ratings are significant ($p<.01$ on Games Howell post hoc tests). The high mean ratings by the native English group when *every* is the object quantifier must contribute to the significant main effect of object QP, as well as the significant interaction between object QP and group. All the patterns found in the native control data are as expected, following the quantification literature.

Moving on to the learner groups, all groups appear to reject inverse scope (i.e., ratings are below 1.5) except for the intermediate EJ group, which has mean ratings indicating acceptance on Types 1b (1.61), 2b (1.76), and 3b (1.92). Games-Howell post hoc tests reveal a number of interesting significant differences between these ratings and the ratings of (i) the native Japanese group, and (ii) the intermediate CJ and KJ groups.[159] The significant differences occur between intermediate EJ ratings and the native Japanese ratings on Types 1b, 2b and 3b ($p<.01$); between the intermediate EJ and the intermediate KJ ratings on Types 2b and 3b ($p<.01$); and between the intermediate EJ and the intermediate CJ ratings on Type 2b ($p<.05$). These between-group differences also contribute to significant effects of group. None of the mean ratings of the other learner groups differ significantly from the native Japanese ratings on any of the inverse scope test types. Nor are there any significant between-learner-group differences that do not involve the intermediate EJ group. However, the mean ratings of all the learner groups are slightly higher than the native Japanese ratings in some cases. In particular, all the learner groups except the advanced KJ group have mean ratings of over 1.0 on Type 1b, while the native Japanese group rating for this type is 0.69; and on Types 2b and 3b the advanced EJ group has mean ratings of over 1.0.

To summarise, subject-wide scope is judged acceptable by all groups on all the relevant test types. Object-wide scope is judged unacceptable in native Japanese and Chinese on all the relevant test types. However, in native English, object-wide

[159] Significant differences also occur between the intermediate EJ ratings and the advanced CJ and KJ ratings (see Appendix 5F). However, these are less interesting, because they are predicted simply on the basis of the advanced groups being significantly more proficient at Japanese than the intermediate groups. Differences between the intermediate groups are not predicted on the basis of proficiency, but rather on the basis of relevant L1 differences.

scope is judged acceptable when the object quantifier is *every*, but unacceptable when the object quantifier is *all*. In L2 Japanese, object-wide scope is judged unacceptable on all the relevant test types by intermediate and advanced learners whose L1 is Chinese or Korean, and by advanced learners whose L1 is English. However, intermediate learners whose L1 is English accept object-wide scope on three of the four relevant types: Type 1b (subj QP = *dareka* 'someone', obj QP = *dono N-mo* 'every N'), Type 2b (subj QP = *dareka* 'someone', obj QP = *subete no-N* 'all the N'), and Type 3b (subj QP = NumP, obj QP = *dono N-mo* 'every N').[160]

5.2.3.3. Results for scrambled QP-QP sentences in native Japanese and L2 Japanese

This section compares the results for the scrambled Japanese QP-QP sentences (Types 1c–d and 3c–d) with those for their non-scrambled counterparts (Types 1a–b and 3a–b). The results for the non-scrambled sentences have already been presented in the previous section. However, for convenience, they are repeated in Table 31 below, along with the results for the scrambled sentences.[161]

[160] Individual results are examined in Section 5.3.4, following.

[161] As with the non-scrambled data (see footnote 156), the native Japanese scores for each of the scrambled test items were checked, prior to calculation of group means, to find out whether any of the mean scores were anomalous, potentially indicating unreliability of a test item. All the scrambled test items were expected to be acceptable, yielding mean scores of close to 3 (+2). This is indeed what was found: all the scrambled test items had mean scores of 2 or higher except items 1.1c (mean = 1.85) and 1.5d (mean = 1.90). These two exceptions still fall into the category of 'acceptance', since they are above the mid-point of 1.5. Thus none of the scrambled test items was considered problematic.

Table 31: Mean ratings and standard deviations for Types 1a–d and 3a–d (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

Group	Type 1a[a] (S>O)	Type 1b[a] (O>S)	Type 1c[a] (S>O)	Type 1d[a] (O>S)	Type 3a[b] (S>O)	Type 3b[b] (O>S)	Type 3c[b] (S>O)	Type 3d[b] (O>S)
	(canonical: SOV)		(scrambled: OSV)		(canonical: SOV)		(scrambled: OSV)	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
JJ (n=20)	2.26 (0.70)	0.69 (0.72)	2.01 (0.71)	2.29 (0.56)	2.86 (0.27)	0.70 (0.62)	2.59 (0.35)	2.43 (0.56)
EJ int (n=18)	2.74 (0.28)	1.61 (0.63)	2.65 (0.35)	2.48 (0.54)	2.81 (0.27)	1.92 (0.77)	2.45 (0.41)	1.96 (0.52)
EJ adv (n=9)	2.82 (0.39)	1.38 (1.12)	2.69 (0.65)	1.64 (1.16)	2.93 (0.14)	1.31 (1.02)	2.76 (0.30)	1.96 (0.85)
CJ int (n=6)	2.07 (0.98)	1.19 (0.77)	2.03 (0.91)	1.83 (1.08)	2.17 (0.85)	0.90 (0.60)	2.27 (0.79)	2.30 (0.43)
CJ adv (n=9)	2.87 (0.17)	1.07 (0.78)	2.64 (0.40)	2.18 (0.72)	2.89 (0.20)	0.82 (0.53)	2.69 (0.23)	1.89 (0.72)
KJ int (n=20)	2.58 (0.45)	1.24 (0.70)	2.33 (0.54)	1.98 (0.88)	2.59 (0.49)	0.92 (0.58)	2.44 (0.42)	2.08 (0.74)
KJ adv (n=15)	2.60 (0.35)	0.64 (0.71)	2.63 (0.42)	1.75 (0.70)	2.85 (0.22)	0.75 (0.59)	2.61 (0.30)	2.13 (0.74)

[a] Types 1a–d: subject = *dareka/mouren/someone*; object = *dono N-mo/mei-CL N/every N*
[b] Types 3a–d: subject = NumP; object = *dono N-mo/mei-CL N/ every N*

Considering the native Japanese data first, the mean ratings on the scrambled sentences are 2.01 (Type 1c) or higher. Thus, both subject-wide and object-wide scope are acceptable in Japanese when the object QP is scrambled over the subject QP. The high native Japanese mean ratings for the scrambled sentences contrast sharply with the low ratings for Types 1b (0.69) and 3b (0.70)—the non-scrambled sentences with object-wide scope interpretations. Object-wide scope is clearly unacceptable in the canonical word order, but acceptable in the scrambled word order. This finding is as expected, following the quantification literature.

A repeated measures ANOVA run on the data in Table 31 confirms that the main effect of scrambling is highly significant ($F(1,90) = 110.68, p<.001$).[162] As in the non-scrambled sentences, the main effects of scope ($F(1,90) = 201.87, p<.001$) and group ($F(6,90) = 3.46, p<.01$) are also significant, but the main effect of subject QP is not ($F(1,90) = 3.12, p=.08$). The scrambling variable interacts significantly with group ($F(6,90) = 6.02, p<.001$), with scope ($F(6,90) = 201.87, p<.001$), with subject QP and group ($F(6,90) = 3.85, p<.01$), and with scope and group ($F(6,90) =$

[162] Levene's test of equality of error variance was significant for Types 1a, 1c, 1d, 3a and 3c (see Appendix 5F). Thus, as in the analysis of the non-scrambled sentences, F-ratios and p-values may not be as reliable as if homogeneity of variance were assured.

8.14, $p<.001$). Other significant interactions are subject QP with group ($F(6,90) = 2.93$, $p<.05$), and scope with group ($F(6,90) = 2.72$, $p<.05$).

Planned within-subjects contrasts were conducted to further quantify each group's differentiation between the scrambled sentences and their non-scrambled, object-wide-scope counterparts. Each group's ratings for Types 1c–d were compared with Type 1b, and Types 3c–d with Type 3b. It was expected that the contrasts would be significant in native Japanese, confirming that object-wide scope on canonical SOV sentences is significantly less acceptable than subject-wide or object-wide scope on scrambled OSV sentences. This is indeed what was found for the native Japanese group ($p<.001$). For the learner groups, the contrasts were significant ($p<.05$) in all but the following cases: intermediate EJ Type 3d with Type 3b ($F=.05$, $p=.83$); advanced EJ Type 1d with Type 1b ($F=2.29$, $p=.17$); intermediate CJ Types 1c and 1d with Type 1b ($F=2.76$, $p=.16$ for Type 1c *v.* Type 1b; $F=1.74$, $p=.24$ for Type 1d *v.* Type 1b), and Type 3c with Type 3d ($F=6.29$, $p=.05$).

These results are relevant to the proposal (Section 5.2.1) that non-native-like lack of differentiation between the canonical object-wide test types (Types 1b and 3b) and the scrambled test types (Types 1c–d and 3c–d) might indicate the following tendencies:

202. a. a tendency to always reject object-wide scope (if low ratings were given to Types 1b and 3b but also (incorrectly) to Types 1d and 3d)
     b. a tendency to always reject inverse scope (if low ratings were given to Types 1b and 3b but also (incorrectly) to Types 1c and 3c)

Since all of the contrasts were significant for the KJ groups and the advanced CJ group, these groups clearly do not fall into either of the categories in (202). The intermediate and advanced EJ groups also do not fall into the categories in (202), because each group's lack of significant differentiation between canonical and scrambled sentences is limited either to the Type 3 sentences with NumP as subject QP (intermediate EJ), or to the Type 1 sentences with *dareka* 'someone' as subject QP (advanced EJ)—that is, it does not occur across both Types 1 and 3. The intermediate CJ group falls into the category in (202b), because the Type 1b/Type 1c contrast and the Type 3b/Type 3c contrast are both non-significant. Thus, this may be evidence of a group tendency to always reject inverse scope. However, even though the differences are non-significant, the intermediate CJ group's mean ratings for Types 1c and 3c are relatively high (>2) and the ratings for Types 1b and 3b are relatively low (<1.2). This response pattern is the same as the native Japanese response pattern, except that the differences are significant in native Japanese. It seems likely, therefore, that rather than indicating a tendency to always reject inverse scope, the lack of significant between-type differences in the intermediate CJ group is due to the small population (n=6) of the group.

To summarise, the L2 group responses to the scrambled test items are similar to the native Japanese responses: both subject-wide and object-wide scope are judged acceptable on scrambled QP-QP sentences. The scrambled sentence ratings are always higher—and, in most cases, significantly higher—than the ratings for object-wide scope on non-scrambled sentences. Consequently, there is no reason to suspect that any group has a tendency to always reject object-wide scope or inverse scope.

5.2.3.4. Results for the native Korean QP-QP task

As described in Section 5.2, the native Korean test sentence design was slightly different to that used in the Japanese test; hence the native Korean results are not compared directly with the other results. The results for the Korean task are presented in Table 32.[163]

Table 32: Mean ratings and standard deviations by 22 native Korean speakers on the Korean QP-QP sentences (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

Type	subj QP/word order/scope (object QP = *motun N* 'all the N')	Mean	SD
K1a	*nwukwunka* 'someone'/non-scrambled/S>O	2.11	0.72
K1b	*nwukwunka* 'someone'/non-scrambled/O>S	0.49	0.44
K1c	*nwukwunka* 'someone'/scrambled/S>O	2.27	0.65
K1d	*nwukwunka* 'someone'/scrambled/O>S	1.24	0.75
K2a	NumP/non-scrambled/S>O	2.35	0.55
K2b	NumP/non-scrambled/O>S	0.98	0.57
K2c	NumP/scrambled/S>O	2.02	1.45
K2d	NumP/scrambled/O>S	0.95	0.88

The subject-wide scope types in Table 32 (Types K1a, K1c, K2a, and K2c) all have mean ratings higher than 2. Thus, subject-wide scope appears to be acceptable in native Korean on both canonically-ordered and scrambled sentences, as in Japanese. By contrast, the ratings for all four object-wide scope types (Types K1b,

[163] Recall from Section 5.2.1 that, although the design of Types K1a–b and K2a–b is the same as Types 2a–b and 4a–b in the Japanese task, the Korean sentences have different vocabulary.

K1d, K2b, and K2d) are below 1.5, indicating that object-wide scope is unacceptable on both canonically-ordered and scrambled sentences. This result is different to what was found for scrambled Japanese QP-QP sentences: object-wide scope was acceptable in the scrambled sentences used in the Japanese task (see Table 31). However, this difference between Japanese and Korean may be due to the different semantic properties of the object quantifiers in the two tasks. Specifically, as noted in Section 5.2.1, the interpretation of *motun* 'all' in Korean appears to be predominantly collective rather than distributive, in contrast to Japanese *dono...mo* 'every'. Distributivity of the object quantifier is required for an object-wide scope reading. Hence, the unacceptability of the object-wide scope test types in the Korean task may follow from the properties of *motun*.

5.2.4. Discussion

The aim of the QP-QP experiment was to investigate two broad predictions based on Full Transfer/Full Access: (i) due to L1 transfer, English-speaking learners of Japanese will exhibit a different developmental path from Korean- and Chinese-speaking learners of Japanese with respect to scope interpretation in Japanese [QP-NOM QP-ACC V] sentences; and (ii) due to access to UG, more advanced English-speaking learners of Japanese will be able to acquire native-like Japanese scope interpretation, despite poverty of the stimulus. These predictions were expressed as the hypotheses in (187) and (188), repeated below in (203) and (204):

203. *QP-QP Hypothesis 1*: (L1 = Korean or Chinese)
Transfer: Lower (and higher) proficiency Korean-speaking and Chinese-speaking learners will reject non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences, regardless of the lexical content of the universally quantified object.

204. a. *QP-QP Hypothesis 2a* (L1 = English)
Transfer: Lower proficiency English-speaking learners will allow non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences when the universally quantified object is *dono N-mo* 'every N' but not when it is *subete no-N* 'all the N'.

b. *QP-QP Hypothesis 2b* (L1 = English)
Access: Higher proficiency English-speaking learners will reject non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences, regardless of the lexical content of the universally quantified object.

This section discusses the results of the experiment with respect to the hypotheses and to the predictions of Full Transfer/Full Access.

The hypotheses in (203) and (204) rest on the premise that object-wide scope is unavailable in Japanese, Chinese and Korean [∃.NOM...∀.ACC...] sentences, but available in English (when the object quantifier is *every*). The native control data confirm that this premise is well-founded. The native Japanese, Chinese and English control results on the non-scrambled QP-QP task (Section 5.2.3.2) provide experimental confirmation of the theoretical claims about scope interpretation in these languages: in English, object-wide scope was acceptable when the object quantifier was *every* and unacceptable when the object quantifier was *all*; in Japanese and Chinese, object-wide scope was unacceptable with both *dono N-mo/mei-CL N* 'every N' and *subete-no N/suoyoude N* 'all the N' as object QP. This corroborates the findings of Phase 2A for Japanese and English (see Chapter 4), and the findings of Lee *et al* (1999a, 1999b) for Chinese and English, using a different test instrument (see Chapter 3).

For Korean, the results of the native Korean QP-QP task (Section 5.2.3.4) provide experimental confirmation of the unacceptability of object-wide scope with the object QP *motun N* 'all the N'. However, this task did not investigate the distributive quantifier *enu N-to* 'every N'. As already noted (footnote 149), pre-experimental data on *enu N-to* 'every N' was collected from two native Korean linguists, who confirmed that object-wide scope is unacceptable on Korean versions of the non-scrambled test sentences (Types 1a–b and 3a–b) with *enu N-to* as the object QP. The theoretical claim that Korean lacks object-wide scope in [S...O...] sentences is thus also supported, although, with respect to the distributive QP *enu N-to* 'every N', support is pre-experimental rather than experimental.

The L2 results from the non-scrambled QP-QP task (Section 5.2.3.2) bear directly on the hypotheses. QP-QP Hypothesis 1 (203) is confirmed. Intermediate and advanced KJ and CJ groups all rejected object-wide scope on all four of the relevant test types (see Table 30). These groups' ratings on the object-wide scope types are generally lower than 1, although there are some exceptions on Type 1b: intermediate CJ, 1.19; advanced CJ, 1.07; intermediate KJ, 1.24. These three higher ratings nonetheless fall below 1.5 and thus meet the criterion of 'rejection'. In addition, none of the KJ or CJ ratings differed significantly from the native Japanese group's ratings on any of the object-wide scope types. Thus the conclusion that QP-QP Hypothesis 1 is confirmed can be justified.

Moving on to the L1 English-speaking learners of Japanese, QP-QP Hypothesis 2a (204a) is partially confirmed: the lower proficiency English-speaking learners accepted non-target-like object-wide scope on Japanese [∃-NOM ∀-ACC V] sentences when the object QP was *dono N-mo* 'every N' (Types 1b and 3b). However, contra the prediction, they also accepted non-target-like object-wide scope on Type 2b, when the object QP was *subete-no N* 'all the N' and the subject QP, *dareka* 'someone'. On these three test types, the group's mean ratings were significantly higher than the native Japanese group's mean ratings. Only on Type 4 (subject QP = NumP, object QP = *subete-no N* 'all the N') did the intermediate EJ learners' mean rating fall just below 1.5 (at 1.32), indicating rejection of object-wide scope. On this type, there was no significant difference between the intermediate EJ rating and the native Japanese rating. The implications of this result (i.e., the lack of clear differentiation between the two object QP types) are discussed below.

The second part of the hypothesis for L1 English-speaking learners of Japanese, QP-QP Hypothesis 2b (204b), is confirmed. The advanced EJ group rejected object-wide scope, with ratings of below 1.5 on all of the four relevant test types (Types 1b, 2b, 3b, and 4b). In fact, this group's ratings on the object-wide scope types are only just below 1.5 (range: 1.04 on Type 4b to 1.38 on Type 1b). Thus they are higher than the ratings by the Chinese and Korean learners. In addition, within-group variation is higher than for any other group, with standard deviations of 1.12 on Type 1b and 1.02 on Type 3b (see Table 30).[164] Nonetheless, the advanced EJ ratings do not differ significantly from the native Japanese ratings on any of the four relevant types, unlike the intermediate EJ ratings.

Although the hypothesis for the intermediate EJ group was not fully confirmed, the overall pattern of the learner results is compatible with the claims of Full Transfer/Full Access. Considering L1 transfer first, the intermediate learner results for Types 1b and 3b (O>S scope, object QP = *dono N-mo* 'every N') are of key importance. The intermediate EJ group accepted object-wide scope on these test types; the intermediate CJ and KJ groups rejected object-wide scope. This difference directly reflects the difference in scope interpretation possibilities between native

[164] The source of this variation is discussed further, below.

English (O>S scope is acceptable) on the one hand and native Chinese and Korean (O>S scope is not acceptable) on the other. Thus, the assumption under Full Transfer/Full Access that the L1 grammar transfers to the interlanguage at the initial state of L2 acquisition readily accounts for the difference between the intermediate EJ group and the intermediate CJ and KJ groups.[165]

The intermediate EJ acceptance of object-wide scope on Type 2b (O>S scope, subject QP = *dareka* 'someone', object QP = *subete-no N* 'all the N') challenges this conclusion. On Type 4b (O>S scope, subject QP = NumP, object QP = *subete-no N* 'all the N'), too, the intermediate EJ group has the highest mean rating (1.32) of all the groups, although this rating falls into the category of rejection (i.e., <1.5). Intermediate EJ learners were predicted to reject object-wide scope when the object QP was *subete-no N* 'all the N' due to transfer of the lexical properties of English *all*, which does not readily allow object-wide scope. However, the fact that, of all three learner groups, the EJ learners are the ones predicted by L1 transfer to allow object-wide scope at all, makes it reasonable to seek an explanation for this result that maintains the L1 transfer account. One explanation could be as follows. At the outset of L2 acquisition of Japanese, the English-Japanese interlanguage contains distinct lexical slots bearing the specific properties of English *every* and *all*, due to transfer from the L1. However, despite the semantic similarities (discussed in Chapter 2) between (i) *every* and *dono...mo* and (ii) *all* and *subete*, the evidence in the Japanese input may be too unclear or infrequent to motivate differentiation in the interlanguage between the two Japanese quantifiers in such a way that *dono...mo* is associated with the slot bearing the properties of English *every*, and *subete*, with the slot bearing the properties of English *all*. This is plausible, considering the subtlety of some of the properties distinguishing *dono N-mo* 'every N' and *subete-no N* 'all the N'. For example, as discussed in Chapter 2, *subete* tends to support a collective interpretation (like English *all*) while *dono...mo* does not (like English *every*).

[165] It is recognised that intermediate learners are no longer at the initial state of L2 acquisition. However, since the scope interpretation contrast between the intermediate EJ group and the intermediate KJ and CJ groups reflects the contrast between native English and native Korean and Chinese, it seems reasonable to propose that the source of the contrast is L1 transfer. Since QP-QP interpretation is a subtle phenomenon, the EJ learners' interlanguage grammars are still influenced by the L1 even at the intermediate level.

However, this distinction is not categorical.[166] Therefore, learners may have no obvious reason to make the hypothesised associations.[167] This may mean that the learners overgeneralise and allow both Japanese quantifiers to have any of the properties of universals transferred from the L1, with the result that, in the intermediate EJ interlanguage, there is no real distinction between *dono...mo* 'every' and *subete* 'all'. Hence, the EJ learners allow both of the Japanese universals to take object-wide scope.

Moving on to the question of access to the mechanisms of UG in L2 acquisition ('Full Access'), it has already been shown that acquisition of the lack of object-wide scope in Japanese is a poverty-of-the-stimulus problem for native speakers of English. The EJ results for Types 1b and 3b indicate that the problem can be overcome, because the advanced EJ group ratings on these test types are lower (i.e., more target-like) than the intermediate EJ group ratings, and, importantly, they indicate rejection of object-wide scope (in contrast to the acceptance by the intermediate EJ group). However, since the advanced EJ group ratings were only just below the mid-point (1.5) on the scale, and the standard deviations were high (Type 1b, mean = 1.38, SD = 1.12; Type 3b, mean = 1.31, SD = 1.02), this could indicate that the participants in this group were answering randomly. Investigation of individual consistency on these two test types sheds light on this issue. Table 33 presents details of individual consistency of response for both EJ groups, and, for comparison, the native Japanese and native English groups. ('Consistent rejection' of a type is defined as ratings of 0 (–2) or 1 (–1) on at least four of the five tokens for that type; 'consistent acceptance', as ratings of 2 (+1) or 3 (+2) on at least four of the

[166] Recall from Chapter 2 that Kawashima (1994a: 136) claims that Japanese *wh*+QPt quantifiers lack a collective interpretation:

i. Dono gakusei-mo kooen de atta.
   every student-QPt park in met
   'Every/each student met in the park'
   (i.e., 'There were submeetings which each student engaged in at the park; *not*, 'The students all met in the park.)

However, not all native speakers of Japanese agree with this judgement: some find the collective interpretation with *dono...mo* is acceptable. When the quantifier is *subete* 'all', on the other hand, the collective interpretation is always readily available.

[167] Indeed, if the interlanguage grammar ultimately becomes native-like with respect to Japanese quantifiers, any direct association of a Japanese quantifier with a lexical slot bearing the properties of English *every* or *all* should eventually be abandoned, since, as discussed in Chapter 2, it seems that that no Japanese quantifier has a direct English equivalent. Thus, eventually, evidence in the input should motivate creation of lexical slots bearing the exact features of the Japanese quantifiers (or, modification of the L1-based lexical slots so that they bear the features of the Japanese quantifiers).

five tokens for that type; and 'inconsistency', as fewer than four positive or four negative ratings on the five tokens for that type.)[168]

Table 33: Consistency of individuals on Types 1b and 3b

group (n)	No. (%) of individuals who demonstrate... 1: consistent rejection on 1b and 3b;	2: consistent acceptance on 1b and 3b;	3: inconsistency on 1b and 3b;	4: consistent rejection on 1b or 3b and inconsistency on the other;	5: consistent acceptance on 1b or 3b and inconsistency on the other;	6: consistent acceptance on 1b or 3b, consistent rejection on the other.
EE (24)	1 (4.17)	11 (45.84)	3 (12.50)	1 (4.17)	5 (20.84)	3 (12.50)
EJ int (18)	2 (11.12)	4 (22.23)	1 (5.56)	3 (16.67)	6 (33.34)	2 (11.12)
EJ adv (9)	6 (66.67)	3 (33.33)	0	0	0	0
JJ (20)	12 (60.00)	0	0	6 (30.00)	0	2 (10.00)

The consistency data for the advanced EJ group in Table 33 dispel any suspicion that the participants answered randomly on Types 1b and 3b: six of the nine advanced EJ participants (66.67%) consistently rejected object-wide scope on both Types 1b and 3b (Category 1), while the remaining three (33.33%) consistently accepted object-wide scope (Category 2). The high standard deviation is thus due to this contrast between those who rejected object-wide scope and those who accepted it. This response pattern is fully compatible with Full Transfer/Full Access. The three advanced EJ learners in Category 2 still have an English-based interlanguage grammar and therefore still accept object-wide scope in Japanese ('Full Transfer'). However, the six advanced EJ learners in Category 1 have a target-like interlanguage grammar and reject object-wide scope. Since acquisition of the lack of object-wide scope is a poverty-of-the-stimulus problem for English-speaking learners of Japanese, the fact that these six learners appear to have overcome poverty of the

[168] No participant consistently selected 'can't decide' for any type.

stimulus implicates the mediation of UG in L2 acquisition ('Full Access').[169]

Consideration of the advanced EJ consistency data within the context of the consistency data of the other three groups presented in Table 33 further supports the Full Transfer/Full Access account. Categories 1 and 4 represent rejection of object-wide scope (Category 1 = consistent rejection on both Types 1b and 3b; Category 4 = consistent rejection on just one of Types 1b or 3b and inconsistency on the other). In the native English group, just one participant falls into each of these categories—8.34% of the group. In the intermediate EJ group, the total rises to 27.79%: two participants in Category 1, three in Category 4. In the advanced EJ group, there are six participants in Category 1, or 66.67% of the group (no advanced EJ participants fall into Category 4). Finally, the native Japanese group has 90% of its participants in either Category 1 (12 participants) or Category 4 (6 participants). A mirror image of this pattern emerges for the categories representing acceptance of object-wide scope (Categories 2 and 5): the majority of the native English group and the intermediate EJ group fall into these categories, compared with only 33.33% of the advanced EJ group (three participants) and none of the native Japanese group. The intermediate EJ response pattern is thus closer to the native English pattern than the native Japanese pattern, but it nonetheless includes some learners who reject object-wide scope in a native-like-way; and the advanced EJ pattern is closer to the native Japanese pattern than the native English pattern, but it includes some learners whose grammar appears to be still characterised by L1-based acceptance of object-wide scope.

[169] To clarify further, the advanced EJ learners' knowledge of the absence of object-wide scope in Japanese SOV QP-QP sentences cannot come from indirect negative evidence in the form of 'noticing' the lack of object-wide scope in a required context because there is no definable required context for inverse object-wide scope. (As explained in Chapter 2, the meaning 'for each book, someone read it' could be expressed in English by (i) *Someone read every book* or by (ii) *Each book was read by someone*. There is no context in which (i) <u>must</u> be used.) In addition, it might be objected that reformulation by a native speaker of non-native-like scope assignment might lead to knowledge of the facts of scope in Japanese. For example, on hearing an L2er utter an SOV QP-QP sentence with a (non-native-like) object-wide scope meaning, a native-Japanese speaker might reformulate the utterance using a scrambled OSV version of the sentence, which allows the object-wide scope meaning. However, all the learner could come to know from this is that it is possible to get an object-wide scope meaning using an OSV sentence. It does not rule out the possibility of an SOV sentence having an object-wide scope meaning. Moreover, based on innumerable discussions of QP-QP scope assignment with (non-linguist) native Japanese speakers, the author is confident that meta-linguistic knowledge of the relevant phenomena—and, consequently, the ability even to talk about these phenomena with ease—is extremely rare. In addition, QP-QP sentences themselves are rare. Therefore, the probability that any of the advanced EJ learners in the experiment had ever discussed, or even formulated, a Japanese QP-QP sentence seems negligibly small. For precisely these reasons, an account of their knowledge of the lack of object-wide scope in terms of mediation by UG, is all the more plausible.

To summarise, the L2 groups' QP-QP results for the non-scrambled sentences generally confirmed the hypotheses. Both intermediate and advanced Chinese- and Korean-speaking learners rejected object-wide scope in Japanese. Intermediate English-speaking learners of Japanese, on the other hand, tended to accept object-wide scope, while advanced English-speaking learners rejected it. The difference between the intermediate EJ learners and the intermediate KJ and CJ learners is readily accounted for in terms of L1 transfer, since it reflects a difference between the L1s: native English allows object-wide scope, while native Korean and native Chinese do not. The fact that the advanced English-speaking learners demonstrated knowledge of the lack of object-wide scope in Japanese SOV sentences is evidence that L2 acquisition is constrained by UG, since acquisition of this knowledge represents a poverty-of-the-stimulus problem for L1 English-speaking learners of Japanese. The data thus support the claims of Full Transfer/Full Access: that both the L1 and UG are sources of knowledge in adult L2 acquisition.

One part of the hypothesis for intermediate English-speaking learners of Japanese was not confirmed: the learners did not differentiate between *dono...mo* 'every' and *subete* 'all', by allowing object-wide scope for the former but not for the latter, despite the fact that L1 transfer predicts such a distinction. Instead, the intermediate EJ learners allowed object-wide scope for *subete* 'all', at least when the subject QP was *dareka* 'someone'. It was suggested that, since the semantic differences between *dono...mo* and *subete* are very subtle, the intermediate EJ learners may not have encountered evidence to motivate association of the former with English 'every' and the latter with English 'all'. Therefore they allowed the L1 property of object-wide scope to apply to both quantifiers. An L1-based account of the intermediate EJ learners' knowledge is thus maintained.

5.3. Main study 2: *Wh*-QP interpretation

This stage of the experimentation uses [*wh*-object ... QP-subject ...] questions to further investigate the predictions of Full Transfer/Full Access. As detailed in Chapter 2, English [*wh*-object ... QP-subject ...] questions and Chinese [QP-subject ... *wh*-object] questions are claimed to allow both individual and pair-list interpretations (e.g., Aoun & Li 1993). In other words (using English for exemplification), a question like *What did everyone buy?* can receive an answer such as *[Everyone bought] apples* (individual answer) or *Jane bought apples, Sam bought* *bananas, Alex bought...* (pair-list answer). By contrast, scrambled Japanese and Korean [*wh*-object ... QP-subject ...] questions are claimed to have only the individual interpretation (e.g., Beck & Kim 1997, Kim 1989 for Korean; Hoji 1985 for Japanese).[170] For English-speaking and Chinese-speaking learners of Japanese, acquisition of the lack of pair-list interpretations in Japanese is a poverty-of-the-stimulus problem: logically, the Japanese input cannot contain direct evidence from which to induce this absence; L1 knowledge will not provide information about the Japanese facts; and classroom instruction does not cover this subtle phenomenon. Thus, Full Transfer/Full Access predicts that, by L1 transfer, L1 English- and L1 Chinese-speaking learners of Japanese will pattern similarly in initially allowing non-native-like pair-list interpretations of Japanese [*wh*-object ... QP-subject ...] questions, while L1 Korean-speaking learners of Japanese will demonstrate native-like rejection of pair-list interpretations from the outset. More advanced L1 English- and L1 Chinese-speaking learners of Japanese may eventually come to know that Japanese lacks the pair-list interpretation via the interaction of subtle indirect evidence in the input (see Chapter 2) with the relevant UG constraints. These predictions are expressed as the hypotheses in (205).

205. a. *Wh-QP Hypothesis 1a* (L1 = English or Chinese)
Lower proficiency L2 learners of Japanese whose L1 is English or Chinese will accept both individual and pair-list interpretations of scrambled Japanese [*wh*-object ... QP-subject ...] questions.

b. *Wh-QP Hypothesis 1b* (L1 = English or Chinese)
Higher proficiency L2 learners of Japanese whose L1 is English or Chinese will differentiate between individual and pair-list interpretations of scrambled Japanese [*wh*-object ... QP-subject ...] questions. They will allow individual interpretations but reject pair-list interpretations.

c. *Wh-QP Hypothesis 2* (L1 = Korean)
Lower and higher proficiency L2 learners of Japanese whose L1 is Korean will differentiate between individual and pair-list interpretations of scrambled Japanese [*wh*-object ... QP-subject ...] questions. They will allow individual interpretations but reject pair-list interpretations.

Native Chinese, English and Korean control groups are investigated, as well as L2 learners of Japanese and native Japanese speakers. The native Korean data

---

[170] As explained in Chapter 2, non-scrambled Japanese [QP-NOM *wh*-ACC V] questions (where the QP is a *wh*+QPt quantifier) are considered to be of marginal acceptability (Hoji 1985; Tomioka 2004). They are not investigated in the experimental part of this dissertation.

yield an unexpected result, namely that pair-list interpretations of scrambled Korean [*wh*-object ... QP-subject ...] questions appear to be acceptable to native speakers of Korean, contra the claims of Beck & Kim (1997) and Kim (1989). In addition, the native Chinese data show that individual interpretations of Chinese [QP-subject ... *wh*-object] questions are not highly acceptable, unlike pair-list interpretations. As will be shown, these findings lead to reformulation of the hypotheses in (205) to predict that (i) L1 Korean-speaking learners of Japanese pattern with L1 English- and L1 Chinese-speaking learners of Japanese in initially allowing non-native-like pair-list interpretations; and (ii) L1 Chinese-speaking learners of Japanese initially reject individual interpretations. With these reformulations, the L2 results support the hypotheses. The following sections present the test design and procedure (Section 5.3.1), the results (Section 5.3.2) and discussion of the results (Section 5.3.3).

5.3.1. Test design and procedure

The test design was very similar to that used for the QP-QP experiment. There were ten participant groups: the six L2 Japanese groups described in Section 5.1.2 (see Table 22), and four native control groups, as detailed in Table 34.[171]

Table 34: Native control participants in the *Wh*-QP task

group	no.	age		details
		range	mean	
JJ	18	19–31	23	University students in Japan.
EE	21	18–23	19	University students in UK.
CC	14	27–35	32	University students resident in UK.
KK	26	21–27	22	University students in Korea.

As in the QP-QP task, the native control participants had all had exposure to languages other than their native language through pre-university and (in most cases) university education in their home countries. None of the Japanese, English or Korean participants were bilingual from early childhood, and none had lived outside their home countries. The mean length of residence in the UK for the native Chinese group was eight months (range: 6–12 months). All but one of the Chinese control participants spoke Mandarin Chinese natively; the remaining one spoke Shanghai

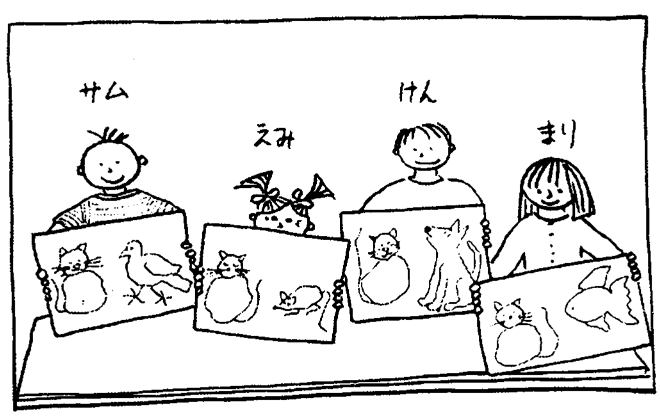
[171] As noted in Section 5.1.2, the 26 native Korean participants in the *Wh*-QP task were the same as those in the QP-QP tasks. Five of the native English participants and one of the native Chinese participants had also completed the QP-QP tasks.

dialect natively. (See Section 5.1.2 for comments on the treatment of Chinese dialects.)

The test instrument was an acceptability judgement task with picture contexts, as in the QP-QP experimentation. In addition to the group variable, one repeated measures variable was manipulated: 'answer type'—individual *v.* pair-list. Each test item comprised a picture with a corresponding question and answer ('Q/A'). There were two answer types representing the two levels of the answer type variable: Type '*Wh*-QPa' consisted of a [*wh*-object...QP-subject...] question with an individual answer; and Type '*Wh*-QPb', a [*wh*-object...QP-subject...] question with a pair-list answer. The structure of the [*wh*-object...QP-subject...] question for both types (*Wh*QPa and *Wh*QPb) is given in (206). The two test types are exemplified in (207).[172]

206. Nani-o daremo-ga V.PAST ka?
what-ACC everyone-NOM V.PAST Q
'What did everyone V? (*scrambled*)'

207. *Wh-QPa and Wh-QPb, example of picture and question:*



(***Translation*** *(L-R):* Sam, Emi, Ken, Mari)

Nani-o daremo-ga kaita no?
what-acc everyone draw.PAST Q
'What did everyone draw?'

a. *Wh-QPa answer (individual answer):*
Neko desu.
cat COP
'A cat.'

[172] For convenience, the Japanese names in the picture in (207) are translated into English below the picture. No English translation was given in the actual test.

b. *Wh-QPb answer (pair-list answer):*[173]
Samu-kun-wa neko to tori-o, Emi-tyan-wa neko to nezumi-o,
Sam-*kun*-TOP cat and bird-ACC Emi-*chan*-TOP cat and mouse-ACC

Ken-kun-wa neko to inu-o, Mari-tyan-wa neko to kingyo-o kaita.
Ken-*kun*-TOP cat and dog-ACC Mari-*chan*-TOP cat and goldfish-ACC drew

'Sam drew a cat and a bird, Emi drew a cat and a mouse, Ken drew a cat and a dog, Mari drew a cat and a goldfish.'

Unlike in the QP-QP experiment, the pictures are not used to manipulate a variable. Instead, they provide a plausible context for both the individual (*Wh*-QPa) and pair-list (*Wh*-QPb) answer types. Thus, if the learners reject the pair-list items (Type *Wh*-QPb) despite the fact that the picture provides a plausible context for those items (i.e., there is no factual mismatch between what is portrayed in the picture and what is described in the *Wh*-QPb answer, as illustrated in (207)), this will provide clear evidence of knowledge of the lack of pair-list readings in Japanese.

Five tokens were created for each type, giving a total of 10 test items. These were mixed randomly with 10 distractor items to create a 20-item test battery. The full test battery is presented in Appendix 6A.

The Japanese task was completed by the six learner groups and the native Japanese group. English, Korean and Chinese versions of the test were created for completion by the native-speaking English, Korean and Chinese control groups. The question forms in Korean and Chinese were as in (208) and (209), respectively, while in English, the form was *What did everyone V?*[174]

208. Mwues-ul nwukwuna-ka V.PAST ni?
what-ACC everyone-NOM V.PAST Q
'What did everyone V? (*scrambled*)'

209. Meigeren V.PAST shenme?
everyone V.PAST what
'What did everyone V?'

[173] *Kun* and *chan* in the gloss are informal male and female titles.
[174] Like Japanese and English, Korean and Chinese also have a choice of words or expressions with the broad meaning of 'everyone'. The use, here, of *nwukwuna* 'everyone' for Korean follows Kim (1989) who treats *nwukwuna* as the equivalent of Japanese *daremo*. Like *daremo*, *nwukwuna* is a *wh*+QPt quantifier: *nwukwu* 'who' + *na* (particle). For Chinese, the choice of *meigeren* 'everyone' follows Aoun & Li (1993), who treat *meigeren* as the equivalent of English *everyone*.

The Chinese question form in (209) omits the floating quantifier *dou* 'all', which is commonly included in the theoretical literature on *wh*-QP interactions in Chinese, as illustrated in (210) taken from Aoun & Li (1993: 59; previously cited in Chapter 2):

210. Meigeren dou maile shenme?
everyone all bought what
'What did everyone buy?'

Dou was omitted because the Japanese question form under investigation (i.e., (206)) does not contain a floating quantifier. Thus, native Chinese data on the question without *dou* should provide more precise information about what could transfer from the L1 to the L1 Chinese-L2 Japanese interlanguage.[175]

The Chinese and Korean test batteries are presented in full in Appendices 6B and 6C, respectively.

As in the QP-QP tasks, there were two presentation orders for the task, Order 2 being the reverse of Order 1. Some participants took the test in Order 1, others in Order 2. This was intended to minimise any potential effect of the order in which the test items appeared.[176]

The distractor items were designed to blend in with the actual test items. They comprised five pairs, each pair sharing a common picture, just as the actual test items comprised five pairs, each pair sharing a common picture. Structurally and lexically, the distractors were similar to the actual test items: two were scrambled *wh*-object questions, beginning with *Nani-o* 'what-ACC' like the actual test tokens, but with non-quantifier subjects (D1x and D5x in Appendix 6A); the other eight were non-scrambled multiple *wh*-questions of the form *Dare-ga nani-o V.PAST* 'Who V.PAST what?'. The most felicitous answer for these multiple *wh*-questions is a pair-list answer (as illustrated in (211), showing item D2x), so the pair-list answers for many of the distractors blended with the pair-list answers of the *Wh*-QPb test items.

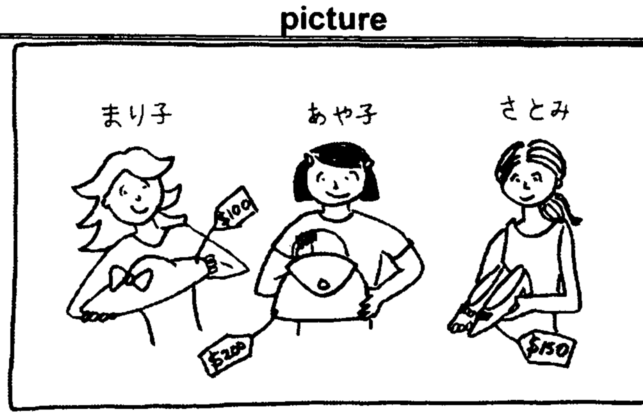
[175] Some native Chinese speakers find the question less natural without *dou*, even though it is grammatical (Yu Jiang, Zhengzheng Wang, personal communication, June 2004). The implications of this for the native Chinese results is discussed, following.
[176] For the native Japanese, native Korean and native Chinese groups, it was not possible to use both test orders, since conditions at the institutions where these tasks were conducted dictated that all participants take the test at the same time. The native Japanese group took Order 2, the native Korean and native Chinese groups, Order 1.

211. *Wh-QP item D2x*

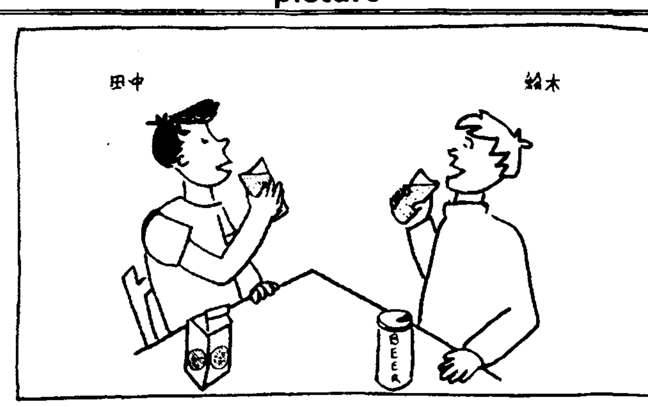
Q/A		picture
Q:	Dare-ga nani-o nonda no? who-NOM what-ACC drank Q 'Who drank what?'	 (*Translation (L-R)*: Tanaka, Suzuki)
A:	Tanaka-san-wa orenzi-zyuusu-o, Tanaka-*san*-TOP orange juice-ACC Suzuki-san-wa biiru-o nonda. Suzuki-*san*-TOP beer-ACC drank. 'Tanaka-*san* drank orange juice and Suzuki-*san* drank beer.'	

Five of the distractors were expected to be judged acceptable, five unacceptable—the same acceptable-to-unacceptable ratio as for the actual test items (in Japanese). The reasons for unacceptability of the distractor items were either that the answer did not match the picture (e.g., D4y shown in (212)), or that it did not fully answer the question (e.g., D2y shown in (213)).

212. *Wh-QP item D4y*

Q/A		picture
Q:	Dare-ga nani-o katta no? who-NOM what-ACC bought Q 'Who bought what?'	 (*Translation (L-R)*: Mariko, Ayako, Satomi)
A:	Daremo-ga bousi-o katta. everyone-NOM hat-ACC bought 'Everyone bought a hat.'	

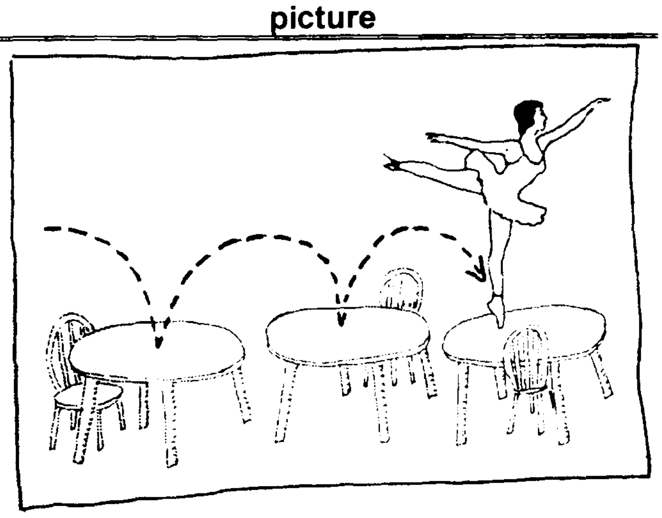
213. *Wh-QP item D2y*

Q/A		picture
Q:	Dare-ga nani-o nonda no? who-NOM what-ACC drank Q 'Who drank what?'	 (*Translation (L-R)*: Tanaka, Suzuki)
A:	Tanaka-san-wa orenzi-zyuusu-o nonda. Tanaka-*san*-TOP orange juice-ACC drank 'Tanaka-*san* drank orange juice.'	

The procedure was as for the QP-QP experiment. Each picture was presented on an overhead projector screen without the corresponding Q/A for 10 seconds, then the Q/A was shown in written form underneath the picture for a further 20 seconds. At the same time as the Q/A was revealed, it was also presented aurally, using an audio tape recorded by a native speaker of the language of the test. Participants rated how possible they found the answer in the context of the question and the picture, using a four-point scale on their answer sheets: –2 = 'definitely not (possible)', –1 = 'not exactly (possible)', 1 = 'kind of (possible)', 2 = 'perfectly (possible)'. A further option of 'can't decide' was also available.

Instructions on how to do the test were given on the test sheet and were also explained orally. The instructions included specific explanation of the rating scale by drawing the participants' attention to the negative ratings corresponding generally to judgements of 'not possible' and the positive ratings corresponding generally to judgements of 'possible'. Four examples were then presented, as illustrated in (214)–(217).

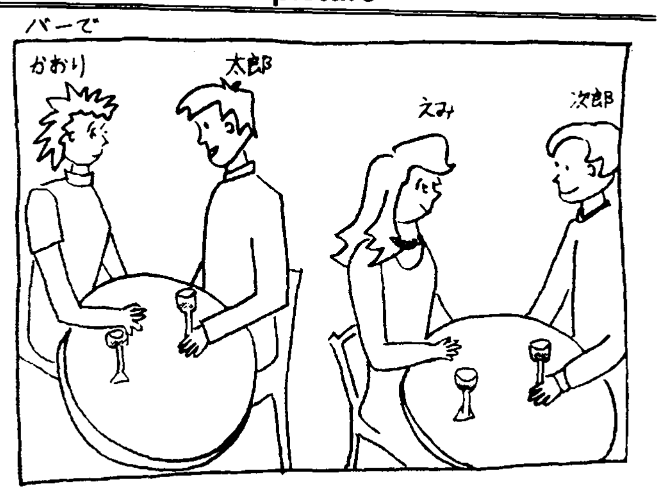
214. *Wh-QP Example 1*

Q/A	picture
Q: Teeburu-no ue-de dare-ga odotta no? table-GEN top-LOC who-NOM danced Q 'Who danced on the table(s)?' A: Barerina desu. ballerina COP 'A ballerina.'	

215. *Wh-QP Example 2*

Q/A	picture
Q: Dare-ga doko-de odotta no? who-NOM where-LOC danced Q 'Who danced where?' A: Barerina-ga iti-dai-no teeburu-no ue-de odotta. ballerina-NOM one-CL-GEN table-GEN top-LOC danced 'A ballerina danced on a table.'	*As Wh-QP Example 1 (above)*

216. *Wh-QP Example 3*

Q/A	picture
Q: Dare-ga dare to osake-o who-NOM who with alcoholic drink-ACC nonda no? drank Q 'Who had a drink with whom?' A: Huta-kumi-no kappuru-ga issyo-ni two-CL-GEN couple-NOM together osake-o nonda. alcoholic drink-ACC drank 'Two couples had a drink together.'	 (*Translation (top line)*: in a bar *(names)*: Kaori, Taro, Emi, Jiro)

217. *Wh-QP Example 4*

Q/A	picture
Q: Dare-ga dare to osake-o nonda no? who-NOM who with alcoholic drink-ACC drank Q 'Who had a drink with whom?' A: Kaori-san-wa Taroo-san to osake-o, Kaori-*san*-TOP Taro-*san* with alcoholic drink-ACC Emi-san-wa Ziroo-san to osake-o nonda. Emi-san-TOP Jiro-*san* with alcoholic drink-ACC drank 'Kaori had a drink with Taro and Emi had a drink with Jiro.'	*As Wh-QP Example 3 (above)*

The examples were presented one by one, with time for discussion of the judgements and for participants to ask questions. They comprised pairs of Q/As sharing a common picture in order to prepare the participants for encountering the same picture with different Q/As in the actual test. Examples 1 and 4 (in (214) and (217), respectively) were intended to exemplify clear cases of 'perfectly possible' answers in the contexts of the corresponding questions and pictures. Examples 2 and 3 (in (215) and (216), respectively) were less straightforward, and ratings varied. It was stressed that there was no 'right' or 'wrong' answer, and participants should try to respond according to their own intuitions.

The test took about 20 minutes to administer, including instructions, with the actual test component lasting 10 minutes.

5.3.2. Results

The *Wh*-QP results are presented in three sections. Details of distractor results are presented first, in Section 5.3.2.1, since these were used to determine whether any participants' data should be excluded from the analysis. Section 5.3.2.2 then presents the native control results, along with analysis of how these results impact on the hypotheses about L2 acquisition of *Wh*-QP scope interpretation. Revised hypotheses are presented at the end of that section. Finally, the L2 results are presented in Section 5.3.2.3.

The data analysis procedure was as in the QP-QP experiment (see Section 5.2.3). The data were coded using Excel. The rating scale of '–2, –1, +1, +2' was transformed to '0, 1, 2, 3'. (Hereafter, the transformed scale is used in this section, with the original scale given in brackets, where relevant.) Responses of 'can't decide' were treated as missing answers and excluded from the analysis, except where 'can't decide' was selected in response to a distractor item, when it was treated as a wrong answer.[177]

The ratings for the five tokens for each test type were averaged for each participant. Mean ratings of below 1.5 (on the transformed scale of 0 to 3) are considered to indicate rejection of the answer; and mean ratings of above 1.5, acceptance.[178] Group means were computed and analysed using SPSS. All statistical details are presented in Appendix 6F.

5.3.2.1. Distractor items

The responses on the distractor items were examined first, to determine whether any participants' data should be excluded due to possible fatigue or lack of understanding of the task. The native Japanese responses to the distractor items, presented as mean ratings (on the transformed scale of 0–3) in Table 35, were taken as the benchmark against which to measure the learner responses.

---

[177] As in the QP-QP experiment, the rate of selection of 'can't decide' was very low. Only 22 instances of 'can't decide' occur among the 3600 responses (=0.61%) included in the test item data presented in Sections 5.3.2.2 and 5.3.2.3 (i.e., the data remaining after exclusions due to distractor results).

[178] Note that, in the test, participants were asked to rate how 'possible' the answers were for each Q/A. In the discussion, ratings indicating judgements of 'possible' (i.e., +1 and +2 on the original rating scale) are taken to indicate 'acceptance' of the answer being judged, and ratings indicating judgements of 'not possible' (i.e., –2 and –1 on the original rating scale) are taken to indicate 'rejection'.

Table 35: Native Japanese ratings on *Wh*-QP distractor items (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

	D1x	D1y	D2x	D2y	D3x	D3y	D4x	D4y	D5x	D5y
Mean	3.00	0.11	3.00	1.90	3.00	1.20	3.00	0.10	0.90	1.80
SD	0.00	0.47	0.00	0.80	0.00	0.80	0.00	0.20	1.10	1.00

It is clear from Table 35 that four distractor items, D2y, D3y, D5x and D5y, yielded somewhat higher levels of variation in rating (SD ≥ 0.8) than the other six. Items D2y (see (213), above), D3y, and D5y were items on which the answer provided in the Q/A was factually true but was pragmatically wrong due to inadequate information. It was expected that the answers on these items would be judged 'impossible' (i.e., ratings of 0 (–2) or 1 (–1)). However, the fact that the answers did not actually conflict with the content of the picture—they merely were not informative enough—is probably the reason for the higher level of variation indicating that some participants gave these answers 'possible' ratings. The high standard deviation (1.10) on item D5x (Q: *What did the boy play?* A: *Paul played the guitar and Laura played the piano*) is likely to have a similar source: in this case, the answer is too informative and therefore pragmatically—but not factually—wrong.

Based on these results, it was decided that distractors D2y, D3y, D5x and D5y were unreliable and should not be considered when investigating individual results for possible cases of wrong answering due to fatigue or other extraneous factors. The criterion for inclusion in the analysis was set at 'correct' responses to at least five of the remaining six distractor items: D1x, D1y, D2x, D3x, D4x, D4y. Correct responses are considered to be ratings of 2 (+1) or 3 (+2) for items D1x, D2x, D3x and D4x, and 0 (–2) or 1 (–1) for items D1y and D4y. None of the native Japanese participants failed to meet this criterion. Among the learner groups, the data from one intermediate EJ participant and one intermediate KJ participant were excluded. Thus data from 19 intermediate EJ participants and 22 intermediate KJ participants are retained for the analysis. The numbers of participants for the other four learner groups remain unchanged from Table 22.

The results for the distractor items on the Chinese, English, and Korean versions of the task were also examined. The results of one native Korean participant were excluded from the analysis due to three 'wrong' answers on the six designated distractor items. Thus data from 25 native Korean participants are retained for the the analysis. The numbers of participants for the native Chinese and English groups remain unchanged from Table 34.

5.3.2.2. Native control data: results, analysis and revised hypotheses

The native control groups' mean ratings (on the scale of 0–3) and standard deviations for Types *Wh*-QPa (individual answers) and *Wh*-QPb (pair-list answers) are presented in Table 36.[179]

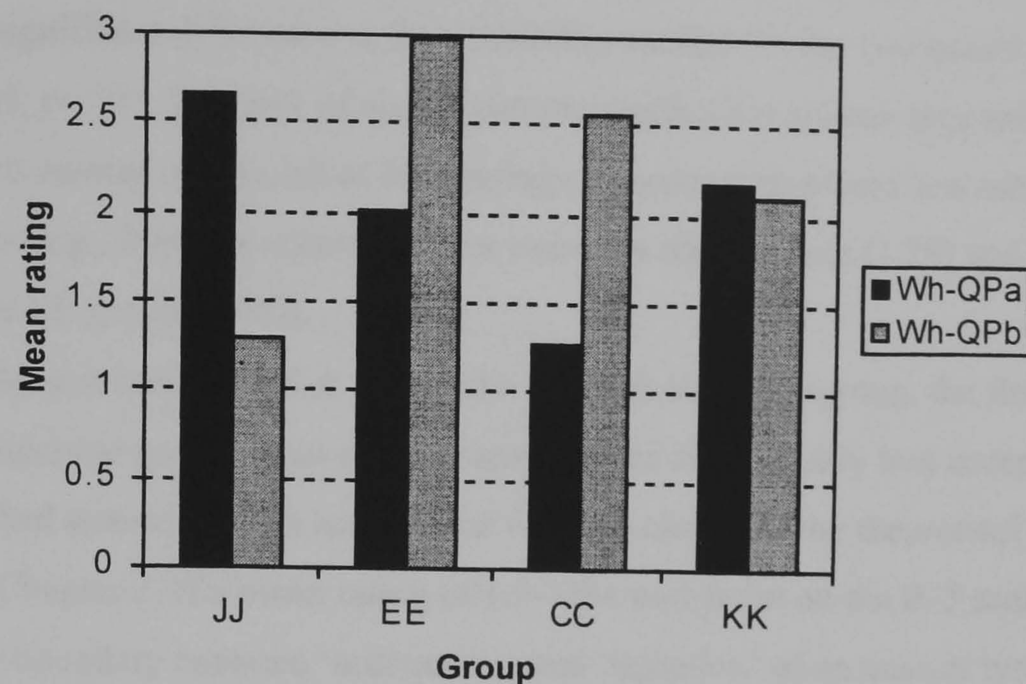
Table 36: Mean ratings on *Wh*-QP task by native control groups (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

	***Wh*-QPa: individual answers**		***Wh*-QPb: pair-list answers**	
Group	Mean	SD	Mean	SD
JJ (n=18)	2.66	0.59	1.29	1.15
EE (n=21)	2.01	0.32	2.98	0.09
CC (n=14)	1.27	0.43	2.57	0.39
KK (n=25)	2.19	0.62	2.10	0.78

The bar chart in Figure 3 displays the mean ratings from Table 36 for each answer type by native group.

[179] Before calculating the group means, the native Japanese mean scores were checked for each test item in order to find out whether any test item had an anomalously high or low value and was therefore potentially unreliable. On Type *Wh*-QPa (individual answers), which were expected to be acceptable, the lowest mean rating was 2.5 on items 2a and 3a. This is close to the highest possible rating of 3 (+2), thus all the *Wh*-QPa items can be considered highly acceptable, as expected. On Type *Wh*-QPb (pair-list answers), which were expected to be unacceptable, the mean ratings ranged from 1.17 (item 1b) to 1.39 (item 2b). These ratings are not as low as they could be. However, the range of 0.22 between the lowest and the highest is very small, thus there was clearly no problem of one *Wh*-QPb item yielding a mean rating that was strikingly higher or lower than the others. All ten Japanese items are thus considered reliable.

Figure 3: Mean ratings for individual (*Wh*-QPa) and pair-list (*Wh*-QPb) answers by native control groups (<1.5 = 'unacceptable'; >1.5 = 'acceptable')



From Table 36 and Figure 3 it is clear that interpretations of [*wh*-object ... QP-subject ...] questions differ according to native language. This observation is confirmed in broad terms by the results of a repeated measures ANOVA (*answer type* x *group*, run on the whole data set, i.e., all ten participant groups) which yielded a highly significant main effect for group ($F(1,150) = 3.8$, $p<.001$), no significant main effect for answer type ($F(1, 150) = 0.870$, $p=.352$), and a highly significant interaction between answer type and group ($F(9, 150) = 7.36$, $p<.001$).[180] These ANOVA results are unsurprising considering the patterns evident in Figure 3 (and without even taking the L2 data into account). The significant main effect for group and significant interaction of group with answer type come from the fact that distinct rating patterns for the two answer types emerge for each group. A paired-samples two-tailed t-test quantifies the within-group differences: native speakers of Japanese find individual answers (*Wh*-QPa) to [*wh*-object ... QP-subject ...] questions

[180] It should be noted that Levene's test run on all the data yields a significant result ($p<.001$) for both *Wh*-QPa and *Wh*-QPb. This means that homogeneity of variance cannot be assumed within the ratings for the two answer types. This is not surprising given the wide range of standard deviations in Table 36 (and Table 37 showing the learner results, following). The implication of this lack of homogeneity of variance is that F-ratios (and consequently p-values) are not as reliable as if homogeneity of variance could be assured. As in the QP-QP experiment, Games-Howell analyses are used for post hoc between-group comparisons, since this procedure is accurate when sample sizes and population variances are unequal.

significantly more possible than pair-list answers (*Wh*-QPb) ($t(17) = 4.36, p<001$): for native speakers of English and native speakers of Chinese it is the opposite: pair-list answers are significantly more possible than individual answers (EE: $t(20) = -13.48, p<.001$; CC: $t(13) = -0.89, p<.001$);[181] and for native speakers of Korean there is no significant difference in the possibility ratings for the two question types ($t(24) = 0.55, p=.71$). The lack of significant main effect for answer type arises because each answer type receives high ratings by some groups and low ratings by other groups (e.g., *Wh*-QPb receives a low rating by the JJ group (1.29) and a high rating by the EE group (2.98)).

Looking in more detail at the results for each language group, the finding for the native Japanese group—that pair-list answers are significantly less acceptable than individual answers—is in accordance with the claims of the theoretical literature outlined in Chapter 2. If a mean rating of 1.5—the mid-point on the 0–3 scale—is taken as the boundary between 'acceptance' and 'rejection' of an answer type (i.e., >1.5 = 'acceptance', <1.5 = 'rejection'), then the Japanese group's mean rating of 1.29 on the *Wh*-QPb condition represents rejection of pair-list readings. Nonetheless, the rating is also high enough to indicate that pair-list readings are not categorically unacceptable in Japanese—categorical unacceptability would be shown by a mean rating of (close to) 0. In addition, the relatively high standard deviation of 1.15 corroborates the frequent observations (e.g., Tomioka 2004) that there is considerable individual variation in native Japanese interpretations of *Wh*-QP questions. Examination of individual results confirms this: six of the 18 JJ participants consistently accepted pair-list readings in Japanese, selecting ratings of 2 (+1) or 3 (+2) on at least four of the five *Wh*-QPb tokens; 10 JJ participants consistently rejected pair-list readings, selecting ratings of 0 (–2) or 1 (–1) on at least four of the five *Wh*-QPb tokens; and two participants were inconsistent in their ratings of *Wh*-QPb tokens, rejecting three and accepting two, or vice versa.

In short, the native Japanese results indicate that (i), in group terms, pair-list readings are significantly less acceptable than individual readings of scrambled

[181] Although individual answers are nonetheless highly acceptable in English (mean rating = 2.01) while they are unacceptable in Chinese (mean rating = 1.27). See discussion, below.

Japanese [*wh*-object ... QP-subject ...] questions; and (ii) a minority of native speakers of Japanese nonetheless find pair-list readings acceptable.[182]

The native English result, whereby both individual and pair-list answers are highly acceptable (mean ratings >2 (+1)), is as expected from the literature cited in Chapter 2 (although the significantly higher acceptability of pair-list readings was not predicted). The native Chinese result, whereby the individual answers are rejected (mean rating = 1.27) but pair-list answers are accepted (mean rating = 2.57), was not expected. Aoun & Li (1993) claimed that both individual and pair-list answers are acceptable in Chinese. This unexpected rejection of individual answers in Chinese may well be explained by the fact, mentioned in Section 5.3.1, that the Chinese question form used did not include the floating quantifier *dou* 'all', unlike the question form discussed by Aoun & Li.[183] The question form used is repeated in (218) for convenience, with the position of the omitted *dou* indicated in brackets.

218. Meigeren (dou) V.PAST shenme?
everyone (all) V.PAST what
'What did everyone V?'

Discussion of (218) with two native speakers of Mandarin Chinese indicates that the most natural answer, with or without *dou*, is a pair-list answer, but if *dou* is stressed, an individual reading becomes more readily available (Yu Jiang, Zhengzheng Wang, personal communication, June 2004). Since *dou* was omitted in the present experiment, it clearly could not be stressed, thus the availability of the individual reading was presumably degraded.

The results for native Korean speakers are also unexpected. Pair-list answers were expected to be unacceptable in native Korean, as in native Japanese. However, the native Korean results indicate that the two answer types are almost equally acceptable, the mean ratings being 2.19 (*Wh*-QPa) and 2.1 (*Wh*-QPb)—just above the 'kind of (possible)' level on the original rating scale, and well above the mid-point of 1.5. A post hoc Games-Howell test reveals that the difference between the native Japanese and native Korean *Wh*-QPb ratings is not actually significant (p=.258).

[182] The assertion for the remainder of this chapter continues to be that Japanese scrambled [*wh*-object...QP-subject...] questions lack a pair-list interpretation. However, it is acknowledged that the native Japanese results testify to a variety of Japanese that allows for such interpretations.
[183] Recall from Section 5.3.1 that *dou* was omitted because the Japanese question form which is the focus of the *Wh*-QP study also does not contain a floating quantifier.

Nonetheless, the finding that there is a highly significant difference between individual and pair-list answers in native Japanese and almost no difference between these two answer types in native Korean contradicts the assertion that Korean behaves in the same way as Japanese with respect to the interpretation of scrambled [*wh*-object ... QP-subject ...] questions. It appears that pair-list interpretations tend to be accepted in native Korean, while they tend to be rejected in native Japanese.[184]

Assuming that these findings about *Wh*-QP interpretation in native English, Chinese and Korean are indeed representative of the three languages,[185] they change the predictions about what transfers to the interlanguage under L1 transfer. For L1 English- and L1 Chinese-speaking learners of Japanese, it was assumed that acceptance of both individual and pair-list readings would transfer to the initial-state interlanguage. For the English-speaking learners this assumption still holds. However, according to the experimental results, the initial-state L1 Chinese-L2 Japanese interlanguage should be characterised by acceptability for pair-list interpretations and unacceptability for individual interpretations. For L1 Korean-speaking learners of Japanese, it was assumed that acceptance of individual interpretations and rejection of pair-list interpretations would transfer to the interlanguage, but the experimental results for native Korean indicate that acceptance of both individual and pair-list readings should transfer. In short, it seems that, for all three groups, acquisition of the lack of pair-list readings in Japanese represents a poverty-of-the-stimulus problem. The hypotheses in (205) are thus abandoned, and the revised hypotheses in (219) (for lower proficiency L2ers) and (220) (for higher proficiency L2ers) are adopted.

219. *Wh-QP Hypothesis 3: Lower proficiency learners*

a. *L1 = Chinese*

Lower proficiency L2 learners of Japanese whose L1 is Chinese will demonstrate non-target-like differentiation between individual and pair-list interpretations of scrambled Japanese [*wh*-object ... QP-subject ...] questions. They will reject individual readings and accept pair-list readings.

b. *L1 = English or Korean*

Lower proficiency L2 learners of Japanese whose L1 is English or Korean will accept both individual and pair-list interpretations of scrambled Japanese [*wh*-object ... QP-subject ...] questions.

---

184 Exploration of why this result occurs is presented in Chapter 6.

185 This assumption is discussed further, at the end of this section.

220. *Wh-QP Hypothesis 4: Higher proficiency learners. L1 = English, Chinese, Korean*
Higher proficiency L2 learners of Japanese whose L1 is English, Chinese or Korean will demonstrate target-like differentiation between individual and pair-list interpretations of scrambled Japanese [*wh*-object ... QP-subject ...] questions. They will accept individual readings and reject pair-list readings.

The interlanguage restructuring that follows from *Hypotheses 3* and *4* (in (219)–(220)) is now most 'drastic' for L1 Chinese-speaking learners of Japanese, since they are predicted to initially reject individual readings of scrambled Japanese [*wh*-object ... QP-subject ...] questions and accept pair-list readings, but eventually accept individual readings and reject pair-list readings. Note, however, that restructuring of the interlanguage with respect to individual readings is not a poverty-of-the-stimulus problem. This is because the target language input could conceivably contain direct evidence for the existence of individual readings in Japanese, in the form of a scrambled [*wh*-object ... QP-subject ...] question in a context that unambiguously requires an individual interpretation. Thus, for the present study, the poverty-of-the-stimulus problem represented by L2 acquisition of the unacceptability of pair-list readings in Japanese remains the primary interest.

The reformulation of the *Wh*-QP hypotheses is, as already pointed out, based on the assumption that the native English, Chinese and Korean data are indeed representative of native knowledge of *Wh*-QP interpretation and are not somehow anomalous. The next section presents evidence from L1 Chinese- and L1 Korean-speaking learners of Japanese that is fully compatible with the findings that (i) individual interpretations are less acceptable than pair-list interpretations in native Chinese; and (ii) pair-list interpretations are acceptable in native Korean. In addition, recent typological evidence about quantifiers in Korean and Japanese (Gill 2002) lends independent support to an account of Korean *nwukwuna* 'everyone' having properties that differ from Japanese *daremo* 'everyone'. Discussion of this typological evidence is pursued further in Chapter 6.

5.3.2.3. L2 Japanese results

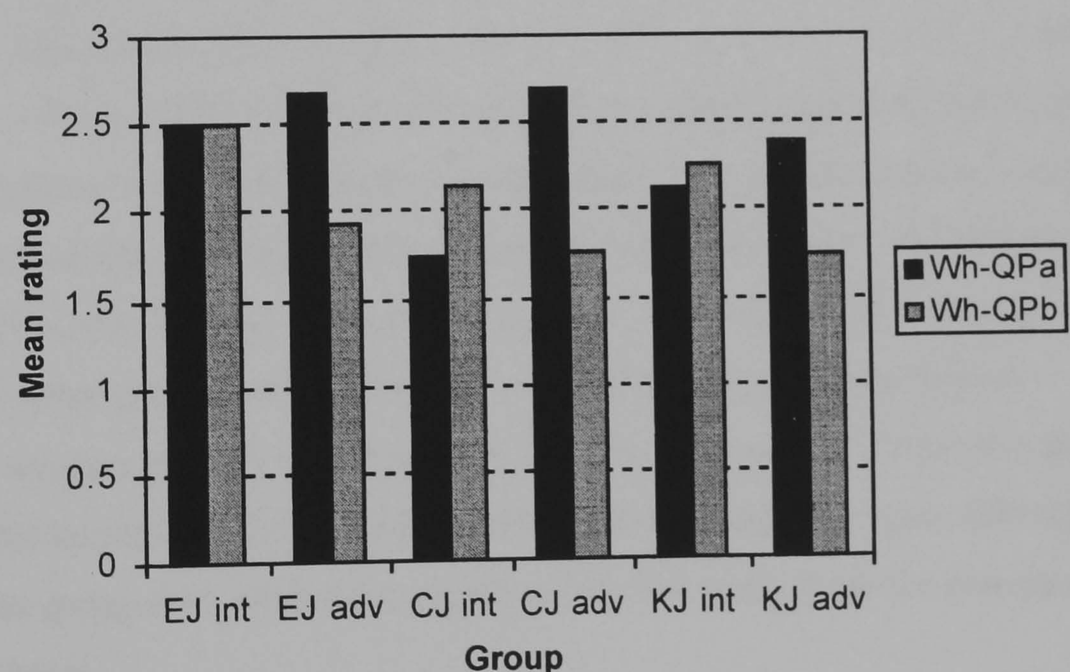
The L2 Japanese groups' mean ratings (on the scale of 0–3) for Types *Wh*-QPa (individual answers) and *Wh*-QPb (pair-list answers) are presented in Table 37.

Table 37: Mean ratings for individual and pair-list answers by learners of Japanese (<1.5 = 'unacceptable'; >1.5 = 'acceptable')

Group	*Wh*-QPa: individual answers		*Wh*-QPb: pair-list answers	
	Mean	SD	Mean	SD
EJ int (n=19)	2.51	0.47	2.51	0.69
EJ adv (n=9)	2.67	0.44	1.93	1.20
CJ int (n=7)	1.74	1.02	2.14	1.06
CJ adv (n=10)	2.70	0.49	1.76	1.25
KJ int (n=22)	2.13	0.84	2.26	0.85
KJ adv (n=15)	2.39	0.78	1.73	1.13

Figure 4 displays the mean ratings from Table 37 for each answer type by learner group.

Figure 4: Mean ratings for individual (*Wh*-QPa) and pair-list (*Wh*-QPb) answer types by learners of Japanese (<1.5 = 'unacceptable'; >1.5 = 'acceptable')



The results of a repeated measures ANOVA (*answer type* x *group*) have already been noted in the previous section (see also full statistical details in Appendix 6F): the main effect of group and the interaction between group and answer type were significant, while the main effect of answer type was not significant. Of importance to the hypotheses under investigation (*Hypotheses 3* and *4* in (219) and (220)) are the within-group differences between the mean ratings for individual (*Wh*-QPa) and pair-list (*Wh*-QPb) answers. Considering the advanced learner groups, first, it is clear from Table 37 and Figure 4 that all three groups have higher mean ratings for individual answers (*Wh*-QPa) than for pair-list answers (*Wh*-QPb), thus exhibiting the same directional differentiation between the two answer types as the native speakers of Japanese (see Table 36 and Figure 3). One-tailed paired samples t-tests reveal that the difference in mean ratings between *Wh*-QPa and *Wh*-QPb is significant only for the advanced CJ group ($t(9) = 1.9, p <.05$). However, the t-test results for the advanced EJ and KJ groups are approaching significance (EJ adv: $t(8) = 1.57, p = .078$; KJ adv: $t(14) = 1.5, p = .079$).

The patterns for the intermediate learner groups differ from the uniform pattern of higher ratings for *Wh*-QPa and lower ratings for *Wh*-QPb demonstrated by the advanced learner groups. For the intermediate EJ group, the two answer types are rated equally acceptable with a mean rating of 2.51 on both *Wh*-QPa and *Wh*-QPb. The Korean and Chinese learners have slightly higher mean ratings on the pair-list answers than the individual answers (KJ int: 2.13 (*Wh*-QPa) *v*. 2.26 (*Wh*-QPb); CJ int: 1.74 (*Wh*-QPa) *v*. 2.14 (*Wh*-QPb)). None of the within-group differences are statistically significant: EJ int: $t(18) = -.36, p = .486$; CJ int: $t(6) = -.53, p = .309$; KJ int: $t(21) = -.47, p = .323$. Post hoc between-group comparisons using the Games Howell procedure also do not reveal any significant differences between learner groups (for example, the intermediate EJ group's mean rating on *Wh*-QPb is not significantly higher than the advanced EJ group's on *Wh*-QPb). In summary, the advanced learner groups all tend towards target-like differentiation between individual answers and pair-list answers, while the intermediate learner groups diverge from target-like differentiation between the two answer types, although each intermediate group does not behave significantly differently from the corresponding advanced group.

Although the direction of the advanced learners' differential acceptance of individual and pair-list answers is the same as that of the native Japanese control group (i.e., higher mean ratings for *Wh*-QPa than *Wh*-QPb), none of the advanced learner groups' mean ratings for pair-list answers fall below 1.5 and thus into the category of 'rejection' of pair-list answers. By contrast, the mean rating of 1.29 by the native Japanese group on *Wh*-QPb was taken to indicate that pair-list answers are rejected in native Japanese. Examination of individual learner consistency on the two test types, presented in Table 38, highlights this difference between the learners and the native Japanese speakers.[186]

Table 38: Consistency of individuals on *Wh*-QPa and *Wh*-QPb Q/A types[a]

Group	***Wh*-QPa** (individual answers) **No. (%) of individuals who demonstrate:**			***Wh*-QPb** (pair-list answers) **No. (%) of individuals who demonstrate:**		
	consistent acceptance	consistent rejection	inconsistency	consistent acceptance	consistent rejection	inconsistency
EJ int (19)	17 (89.5)	1 (5.3)	1 (5.3)	16 (84.2)	2 (10.5)	1 (5.3)
EJ adv (9)	8 (88.9)	0	1 (11.1)	5 (55.6)	4 (44.4)	0
CJ int (7)	2 (28.6)	2 (28.6)	3 (42.8)	4 (57.1)	1 (14.3)	2 (28.6)
CJ adv (10)	9 (90.0)	0	1 (10.0)	6 (60.0)	4 (40.0)	0
KJ int (22)	15 (68.2)	4 (14.1)	3 (17.7)	15 (68.2)	3 (13.6)	4 (14.1)
KJ adv (15)	11 (73.3)	1 (6.7)	3 (20.0)	7 (46.7)	6 (40.0)	2 (13.3)
JJ (18)	17 (94.4)	1 (5.6)	0	6 (33.3)	10 (55.6)	2 (11.1)

[a] *Criteria for consistency*:
'Consistent acceptance' = selection of 2 (+1) or 3 (+2) on at least four of the five test tokens.
'Consistent rejection' = selection of 0 (–2) or 1 (–1) on at least four of the five test tokens.
'Inconsistency' = ratings corresponding neither to consistent acceptance nor consistent rejection.

As Table 38 shows, over half of the learners in all three advanced learner groups (as well as in all three intermediate learner groups) consistently accept pair-list answers in the *Wh*-QPb condition. In the native Japanese group, on the other hand, over half of the participants consistently reject pair-list answers. Table 38 also serves to show that individual learners were generally consistent in their acceptance or rejection of each answer type. The highest rates of inconsistency occur in the smallest group: three of the seven intermediate CJ learners (42.8%) demonstrate inconsistency on the *Wh*-QPa condition and two (28.6%) on the *Wh*-QPb condition. Thus the high standard deviations occurring for some ratings in Table 37 are due mainly to some participants consistently accepting the tokens of that answer type while others consistently reject them.

To summarise, the advanced learner groups demonstrate more target-like behaviour than the intermediate groups with respect to the key target language phenomenon in this part of the study: the acquisition of the lack of pair-list answers in scrambled Japanese [*wh*-object ... QP-subject ...] questions. For the advanced

[186] No individual participant demonstrated consistent selection of 'can't decide'.

learners, as for the native Japanese speakers, pair-list answers are less acceptable than individual answers; while for the intermediate learners, there is no target-like differentiation between the two answer types. Specifically, the proportion of participants in each advanced learner group who consistently reject pair-list interpretations in Japanese is greater than in each intermediate learner group.

5.3.3. Discussion

The revised hypotheses in Section 5.3.2.2 (219–220) made predictions about learner development with respect to both individual and pair-list answers to Japanese [*wh*-object...QP-subject...] questions. These predictions are schematised in (221):

221.

		individual answers:	*pair-list answers:*
a.	*CJ intermediate:* (Hypothesis 3a)	reject	accept
	CJ advanced: (Hypothesis 4)	accept	reject
b.	*EJ & KJ intermediate:* (Hypothesis 3b)	accept	accept
	EJ & KJ advanced: (Hypothesis 4)	accept	reject

The advanced learners of all three L1s are predicted to achieve target-like acceptance of individual answers and rejection of pair-list answers. However, the 'starting point' for the L1 Chinese-speaking learners differs from that of the L1 Korean- and L1 English-speaking learners, as (221) shows: intermediate Chinese-speaking learners are predicted to reject individual answers and accept pair-list answers; and intermediate English-speaking and Korean-speaking learners are predicted to accept both answer types equally.[187]

Although both answer types are investigated, it is important to note that the status of individual answers is not the same as that of pair-list answers in the context

[187] Again, it is acknowledged that the interlanguage grammar of intermediate learners does not represent the initial state of L2 acquisition. However, since *Wh*-QP interpretation is a subtle phenomenon, the assumption behind the predictions is that the intermediate interlanguage grammars will still reflect the L1 grammars.

of the present research. Specifically, as explained in Section 5.3.2.2, acquisition of the lack of pair-list answers in Japanese is a poverty-of-the-stimulus problem for all three learner groups; however, acquisition of the availability of individual answers is not. This is because the Japanese input could conceivably contain [*wh*-object...QP-subject...] questions in unambiguous individual-answer contexts. These would constitute positive evidence from which the availability of individual answers could be induced. Consequently, successful acquisition of the availability of individual answers in Japanese is not informative about the role of UG in L2 acquisition (even in the case of Chinese-speaking learners, for whom L1 transfer predicts initial rejection of individual answers). However, for acquisition of the lack of pair-list answers in Japanese, there can be no direct evidence in the input. Thus evidence of target-like behaviour with respect to pair-list answers would implicate mediation by UG in the L2 acquisition process. Hence, pair-list answers are of key interest to the questions motivating this research.

With this in mind, the advanced learner results are considered first. The relevant hypothesis is Hypothesis 4 (220), which predicted that advanced learners of all three L1 groups would demonstrate target-like differentiation between individual and pair-list answers, accepting individual answers and rejecting pair-list answers. This hypothesis is partially confirmed: all three advanced learner groups show the same directional differentiation between individual answers and pair-list answers as the native Japanese group, in that their ratings for individual answers are considerably higher than for pair-list answers. The difference between ratings for individual answers and for pair-list answers is statistically significant for the advanced Chinese-speaking learners and close to significant ($p<.08$) for the advanced English-speaking and Korean-speaking learners. However, none of the advanced learner groups have mean ratings for pair-list answers that are below 1.5 and hence fall into the category of 'rejection', unlike the native Japanese group.

This means that, in group terms, the advanced learners have not acquired the lack of pair-list interpretation of Japanese [*wh*-object...QP-subject...] questions. However, as seen in Table 38, a number of advanced learners in each L1 group demonstrated target-like consistent rejection of pair-list answers in Japanese. These individual results show that the target syntax-semantics architecture can be acquired, despite poverty of the stimulus. In other words, they provide further support for the claim that L2 acquisition is constrained by UG, since, for each L1 background, knowledge of the lack of pair-list readings in Japanese is underdetermined by L1 knowledge and by direct evidence in the input. The fact that fewer than half of the learners in each advanced group demonstrate consistent knowledge of the lack of pair-list readings in Japanese can be ascribed to the indirect evidence to motivate the relevant interlanguage structuring being very sparse. Thus, those who have not acquired the target-like grammar have presumably not yet assimilated (enough of) the relevant evidence.

The hypotheses about the intermediate learners are relevant to the question of L1 transfer. The hypothesis about the L1 English-speaking and L1 Korean-speaking learners (Hypothesis 3b (219b)) is learners is confirmed: they accepted both individual and pair-list answers, as predicted by L1 transfer. For the L1 Chinese-speaking learners, Hypothesis 3a (219a) predicted rejection of individual answers and acceptance of pair-list answers. The results show that the Chinese-speaking learners in fact accepted both. However, the mean rating for individual answers by this group was lower than for pair-list answers, and it was also lower (at 1.74) than the mean ratings for individual answers by the intermediate L1 English- and L1 Korean-speaking learners (≥2.13). Thus, although the L1 Chinese group did not reject individual answers, the group's rating for this answer type was nonetheless considerably lower than the ratings by the other two groups. The pattern of this between-groups difference is exactly as expected based on L1 transfer, even though it is not as striking as the difference between individual answers in native Chinese (mean rating, 1.27) and in native English and Korean (≥2.19). In this context, it is important to note that, as pointed out above, acquisition of the target high acceptability of individual answers in Japanese is not a poverty-of-the-stimulus problem: positive evidence for induction of this acceptability may be available in the input. In addition, learners' interlanguage at the intermediate level does not represent the initial state of L2 acquisition. The L1 Chinese-speaking intermediate learners thus may have already encountered enough evidence to motivate positive-evidence-based restructuring of their interlanguage with respect to individual answers. Since the acquisition of target-like high acceptance of individual answers is not a poverty-of-the-stimulus problem, it is not unexpected that this might proceed at a different rate to acquisition of the lack of pair-list answers.

In short, the predictions of the hypotheses are largely supported, and the overall pattern of the L2 results is compatible with an account in terms of Full

Transfer/Full Access. For pair-list answers, the results support (i) an L1 transfer-based prediction that pair-list answers will be highly acceptable to lower proficiency learners; and (ii) a UG-access-based prediction that advanced learners will demonstrate knowledge of the unacceptability of pair-list answers in Japanese. [188]

An additional plausible implication of the L2 results concerns the reliability of the native control data. Specifically, the intermediate CJ and KJ results lend credence to the reliability of the unexpected findings for native Chinese and native Korean detailed in Section 5.3.2.2—namely that the acceptability of individual answers in native Chinese is low, and the acceptability of pair-list answers in native Korean is high. These two patterns are replicated in the intermediate CJ and KJ results, respectively: the intermediate L1 Chinese-speaking learners of Japanese have lower mean ratings for individual answers than pair-list answers (1.74 *v.* 2.14), like the native Chinese group (1.27 *v.* 2.57) (although the difference is significant only for the native Chinese group); and the intermediate L1 Korean-speaking learners of Japanese have approximately equally high levels of acceptance for both individual answers and pair-list answers (2.13 *v.* 2.16), like the native Korean group (2.19 *v.* 2.1). This replication of the native patterns in the intermediate L2 data is as predicted by L1 transfer. Although replication of the native patterns in the L2 data does not suffice as proof of the reliability of the native data, if the native Chinese and native Korean data had been randomly anomalous, it is unlikely that the same randomly anomalous patterns would be reproduced in the intermediate CJ and KJ data.

Both the native Korean data and the L2 data have implications for theories of pair-list interpretations of *wh*-QP questions. The key findings to account for are (i) the discovery that pair-list interpretations appear to be highly acceptable in native Korean, unlike in native Japanese; and (ii) the discovery that, while L2 acquisition of the relative unacceptability of pair-list readings in Japanese is possible despite poverty of the stimulus, the indirect evidence motivating the acquisition must be very subtle since fewer than half of the advanced learners in each learner group were able to acquire the phenomenon. Reappraisal of the theories of *wh*-QP interpretation in light of these discoveries is explored in the next chapter. In addition, the finding that

[188] An outcome incompatible with Full Transfer/Full Access would have been, for example, if the L1 English-speaking learners of Japanese had rejected individual answers.

a minority of native Japanese-speakers accepted pair-list interpretations is also discussed further in Chapter 6.

5.4. Conclusion

The broader research questions ((182) and (183)) behind the hypotheses of the QP-QP task and the *Wh*-QP task are repeated in (222) and (223):

222. Do L2ers show divergence with respect to a Target language phenomenon *P* when their L1s are typologically distinct with respect to *P*?

223. When *P* represents an L2 poverty-of-the-stimulus problem, are L2ers able to overcome the problem and acquire *P*?

The results support affirmative answers to both questions. In the QP-QP task, the behaviour of English-speaking learners of Japanese diverged from that of Chinese-speaking and Korean-speaking learners in a way that reflected the divergent L1s with respect to QP scope interpretation. In both the QP-QP task and the *Wh*-QP task, the advanced learner results provided evidence of emerging target-like knowledge of scope interpretation in doubly quantified sentences and [*wh*-object...QP-subject...] questions, despite poverty of the stimulus. These results indicate that (i) the L1 is a privileged source of knowledge in L2 acquisition; and (ii) the mechanisms of UG are available in L2 acquisition, enabling acquisition under poverty of the stimulus. As such, the results lend support to the Full Transfer/Full Access theory of L2 acquisition.

6. Implications and outstanding issues

6.0. Introduction

The main claims arising from the present research project are those pertaining to L2 acquisition theory detailed in the previous chapter. Specifically, the results of the experimental investigation of quantifier scope interpretation by adult L2 learners of Japanese were shown to support the Full Transfer/Full Access model of L2 acquisition. In other words, they support the two-part conclusion that (i) L1 knowledge transfers to the interlanguage grammar at the start of L2 acquisition; and (ii) L2 acquisition is guided by UG.

In addition to supporting this conclusion, the results give rise to a number of issues for further exploration. Discussion of these outstanding issues is presented in this chapter. Section 6.1 reviews the findings about quantifier scope interpretation in native Japanese, Chinese, Korean, and English in the context of (i) the theoretical claims about quantifier scope interpretation in these languages, and (ii) the findings of other studies of quantifier scope interpretation. The unexpected native Korean acceptance of pair-list interpretations of scrambled Korean [*wh*-object...QP-subject...] questions is discussed in this section. Section 6.2 explores the implications of the L2 results for the theories of quantifier scope interpretation outlined in Chapter 2. Section 6.3 briefly comments on the experimental methodology and discusses its strengths and limitations. Finally, Section 6.4 summarises the chapter.

6.1. Implications of the native experimental data

A significant contribution of the present research—in addition to investigating L2 acquisition—is to provide quantitative experimental data on quantifier scope interpretation in native Japanese, English, Chinese and Korean. Although some relevant data are already available from existing studies, the results of the present research considerably augment this database.[189] Prior to the present study, data were particularly lacking on the interpretation of QP-QP sentences in native Japanese and

[189] As reported in Chapter 3, the full-scale experimental studies of Kurtzman & MacDonald (1993) and Lee *et al.* (1999b) investigate scope interpretation in QP-QP sentences in native English and Chinese. In addition, Sano (2004), reports on the judgements of 10 native speakers of Japanese about one Japanese QP-QP sentence. For *wh*-QP interpretation, Miyamoto & Yamane (1996) report on the judgements of five native speakers of English. To the author's knowledge, these are the only experimental studies of quantifier scope interpretation in the languages under investigation.

native Korean, and on the interpretation of *wh*-QP questions in all four languages. The aim of the present section is to review the findings of the main studies for the native control groups, and to explore the implications of these findings. Section 6.1.1 focuses on the QP-QP study, and Section 6.1.2, on the *wh*-QP study.

6.1.1. QP-QP study: native experimental data

This section is organised as follows. Section 6.1.1.1 reviews the native Japanese results and discusses their implications for the theories of quantifier scope interpretation in Japanese, presented in Chapter 2. Section 6.1.1.2 reviews the native Korean results. Section 6.1.1.3 reviews the native English and native Chinese results and compares them with the results of the previous studies by Kurtzman & MacDonald (1993) (English) and Lee *et al.* (1999b) (Chinese and English).

6.1.1.1. Native Japanese data

Judgements were sought on five tokens each of the twelve test types summarised in Table 39.

Table 39: Test types, QP-QP experiment

type	sentence structure	scope
1a:	*dareka* 'someone'-NOM *dono N-mo* 'every N'-ACC V.PAST	S>O
1b	*dareka* 'someone'-NOM *dono N-mo* 'every N'-ACC V.PAST	O>S
1c	*dono N-mo* 'every N'-ACC *dareka* 'someone'-NOM V.PAST	S>O
1d	*dono N-mo* 'every N'-ACC *dareka* 'someone'-NOM V.PAST	O>S
2a	*dareka* 'someone'-NOM *dono N-mo* 'every N'-ACC V.PAST	S>O
2b	*dareka* 'someone'-NOM *dono N-mo* 'every N'-ACC V.PAST	O>S
3a	NumP-NOM *subete-no N* 'all the N'-ACC V.PAST	S>O
3b	NumP-NOM *subete-no N* 'all the N'-ACC V.PAST	O>S
3c	*subete-no N* 'all the N'-ACC NumP-NOM V.PAST	S>O
3d	*subete-no N* 'all the N'-ACC NumP-NOM V.PAST	O>S
4a	NumP-NOM *subete-no N* 'all the N'-ACC V.PAST	S>O
4b	NumP-NOM *subete-no N* 'all the N'-ACC V.PAST	O>S

The verb in each of the types in Table 39 was an active, transitive verb in the past tense. The scope variable was manipulated by means of pictures depicting either the subject-wide or the object-wide interpretation of the sentence. Twenty native speakers of Japanese judged the test sentences. (See Chapter 5 for details of the procedure; see Appendix 5 for the full set of test items and pictures.)

As reported in detail in Chapter 5, the 'a', 'c' and 'd' test types (SOV order with S>O scope, OSV order with S>O scope, and OSV order with O>S scope, respectively) were judged acceptable with group mean ratings of at least 2.01 on a scale of 0 to 3, where 0 indicated outright rejection and 3 indicated complete acceptance.[190] This is as predicted according to the literature on QP-QP scope in Japanese (Hayashishita 1999, 2000a, 2000b; Hoji 1985; Kuno & Takami 2002; Kuroda 1970). However, there is disagreement in the theoretical literature regarding the acceptability of SOV QP-QP sentences with object-wide scope, like the 'b' types in Table 39. To restate the debate in brief (see Chapter 2 for details), Kuroda (1970) and Hoji (1985) claim that object-wide scope is not possible in Japanese SOV QP-QP sentences, while Kuno & Takami (2002) and Hayashishita (1999a, 1999b, 2000) claim that it is. Kuno & Takami (2002) assert that when the subject QP is an existential and the object QP a universal, as in all the test types in Table 39, both subject-wide and object-wide scope are possible in Japanese.[191] Hayashishita (1999a, 1999b, 2000) claims that object-wide scope is available cross-linguistically in [S…O…] sentences, as long as the object QP can denote a specific group or individual. Thus, object-wide scope should be available in all of the 'b' sentences in Table 39, according to Hayashishita, because both *dono N-mo* 'every N' and *subete-no N* 'all the N' can denote a specific group.[192] In the context of this debate, the results of the present study thus support Kuroda (1970) and Hoji (1985): each of the four 'b' test types—SOV order with object-wide scope—was judged unacceptable, with group mean ratings of 0.73 or lower.

Examination of the individual response patterns further illustrates the lack of acceptability of object-wide scope across the four 'b' test types, as shown in Table 40.

---

[190] Recall that in the actual test, the four-point scale was –2, –1, +1, +2. This was transformed to a scale of 0 to 3 for the analysis. The transformed scale is used throughout this section, with the original scale ratings in brackets where relevant.
[191] On QP-QP sentences with a universal subject and existential object, Kuno & Takami's (2002) claim is in accordance with that of Hoji (1985): that object-wide scope is unavailable.
[192] An example of a non-group-denoting QP is *kanari-no-kazu-no N* 'a good number of N'. In the discussion of Hayashishita's analysis, it was suggested that augmentation of this analysis to include Tomioka's (2004) proposal about the anti-topic status of Japanese *wh*+QPt quantifiers (i.e., quantifiers comprising a *wh*-word and a quantificational particle [QPt], such as *dono N-mo* 'every N, *lit.* 'which N-QPt') may account for the lack of object-wide scope in Japanese SOV QP-QP sentences with a *wh*+QPt quantifier as object (see Chapter 2). However, without augmentation, Hayashishita (1999, 2000a, 2000b) does not account for any lack of object-wide scope in Japanese QP-QP sentences when the object QP is a group-denoting QP.

Table 40: Summary of consistency of individual native Japanese response patterns on the four SOV object-wide scope types (Types 1b, 2b, 3b and 4b)[a]

participant[b] (n = 20)	type 1b	type 2b	type 3b	type 4b
subject QP:	*dareka* 'someone'	*dareka* 'someone'	NumP	NumP
object QP:	*dono N-mo* 'every N'	*subete-no N* 'all the N'	*dono N-mo* 'every N'	*subete-no N* 'all the N'
JJ19*	rejected	rejected	rejected	rejected
JJ20	rejected	inconsistent	inconsistent	*ACCEPTED*
JJ22*	rejected	rejected	rejected	rejected
JJ23	rejected	inconsistent	rejected	rejected
JJ24*	rejected	rejected	rejected	rejected
JJ25	rejected	rejected	*ACCEPTED*	inconsistent
JJ26	rejected	rejected	rejected	inconsistent
JJ27*	rejected	rejected	rejected	rejected
JJ28	inconsistent	*ACCEPTED*	rejected	inconsistent
JJ29*	rejected	rejected	rejected	rejected
JJ30	rejected	rejected	inconsistent	inconsistent
JJ31*	rejected	rejected	rejected	rejected
JJ32*	rejected	rejected	rejected	rejected
JJ33	inconsistent	rejected	rejected	rejected
JJ34*	rejected	rejected	rejected	rejected
JJ35	inconsistent	rejected	rejected	rejected
JJ36	*ACCEPTED*	rejected	rejected	inconsistent
JJ38*	rejected	rejected	rejected	rejected
JJ39*	rejected	rejected	rejected	rejected
JJ40	rejected	inconsistent	inconsistent	rejected
Total rejected:	16 (80%)	16 (80%)	16 (80%)	14 (70%)
Total accepted:	1 (5%)	1 (5%)	1 (5%)	1 (5%)
Total inconsistent:	3 (15%)	3 (15%)	3 (15%)	5 (25%)

[a] 'rejected' = 'consistent rejection': selection of ratings of 0 (–2) or 1 (–1) on at least four of the five tokens for that type.
'accepted' = 'consistent acceptance': selection of ratings of 2 (+1) or 3 (+2) on at least four of the five tokens for that type.
'inconsistent' = ratings corresponding neither to consistent acceptance nor to consistent rejection.
[b] Asterisks indicate those participants who consistently rejected object-wide scope on all four test types.

The 'Total rejected' line in Table 40 shows that, for each test type, there were at least 14 individuals who consistently rejected object-wide scope on that type, while the 'Total accepted' line shows there was a maximum of one who consistently accepted object-wide scope on each type (the notions of 'consistent rejection' and 'consistent acceptance' are used here with the definitions given in Note 'a' in Table 40). Each instance of consistent acceptance of object-wide scope is by a different individual: none of the 20 native Japanese participants consistently accepted object-wide scope across all four of the relevant test types.

The reliability of this result of rejection of object-wide scope in Japanese SOV QP-QP sentences is confirmed—albeit on a small scale—by Sano's (2004) finding that 8 out of 10 native Japanese speakers rejected object-wide scope on one sentence with the same structure as Type 1b (subject = *dareka* 'someone', object = *dono N-mo* 'every N'). Similarly, the pilot results of the present study (Chapter 4) also yielded results showing native Japanese rejection of object-wide scope in SOV QP-QP sentences of the types in Table 40.

Given the individual response patterns in Table 40 along with the group mean ratings indicating group rejection of object-wide scope on SOV QP-QP sentences, the data from the present study do not provide support for theories claiming that object-wide scope is available in Japanese SOV QP-QP sentences (i.e., the accounts of Kuno & Takami 2002 and Hayashishita 1999, 2000a, 2000b). Under both Kuno & Takami (2002) and Hayashishita (1999, 2000a, 2000b), pragmatic factors may account for failure of an object-wide scope interpretation to arise (see Chapter 2 for details). The pragmatic solution is expressed differently in each account: under Kuno & Takami (2002), object-wide scope is ruled out by the lack of a discourse context favouring such an interpretation; under Hayashishita (1999, 2000a, 2000b) it may be ruled out by the lack of a group referent in the discourse for the (group-denoting) object QP. However, such pragmatic accounts of the lack of acceptance of object-wide scope in the SOV QP-QP sentences do not seem viable in the present experiment. The picture contexts used for each test token were designed to provide a pragmatically plausible context (including an appropriate group referent) for the intended scope interpretation of that token. The task required the participants to judge, for each token, whether the sentence matched the picture. The same pictures were used for the object-wide scope conditions in both the SOV types (the 'b' types) and the OSV (scrambled) types (the 'd' types). The fact that the participants accepted object-wide scope on the scrambled types (mean rating $\geq$ 2.29 on the scale of 0–3) shows that the pictures were successful: if the picture contexts had been pragmatically implausible, the acceptability of object-wide scope on the scrambled test types would have been severely degraded.

In short, the native Japanese results for the QP-QP sentence types investigated here are highly compatible with claims that object-wide scope is not readily available in Japanese SOV doubly-quantified sentences (Kuroda 1970; Hoji 1985); and they do not support analyses claiming that Japanese QP-QP sentences are essentially ambiguous (Kuno & Takami 2002; Hayashishita 1999, 2000a, 2000b). However, the present results do not say anything about QP-QP sentences with verbs other than active, transitive verbs in the simple past; nor do they inform us about quantifier combinations other than those used here. In addition, it is acknowledged that the data in Table 40 reveal some inconsistency in the individual response patterns: although 10 participants (50%) consistently rejected object-wide scope on all four of the SOV test types, the remaining 10 consistently rejected object-wide scope on only three (four participants), two (two participants), or one (two participants) of the four test types, and they are inconsistent, or, in four instances, they demonstrate acceptance, on the others. This degree of inconsistency remaining in the results, despite rigorous refining of the test instrument (see Chapters 4 and 5), may be a quantitative reflection of the frequently-alluded-to subtlety and fragility, if not outright difficulty, of judgements about quantifier scope interpretation. As such, some inconsistency may be inevitable in studies of quantifier scope interpretation.

6.1.1.2. Native Korean data

Korean, like Japanese, is claimed to exhibit scope rigidity in SOV QP-QP sentences, and scope ambiguity in scrambled OSV QP-QP sentences. This claim is supported in Kim (1989), Kim & Larson (1989) and Sohn (1995) using sentences with *nwukwunka* 'someone' as the subject QP and *nwukwuna* 'everyone' as the object QP, and an active transitive verb, as shown in (22) and (225) (Kim 1989: 366 (24a), 367 (28a), previously cited in Chapter 2).

224. Nwukwunka-ka nwukwuna-lul chodayhatta. unambiguous: S>O; *O>S
someone-NOM everyone-ACC invited
'Someone invited everyone.'

225. Nwukwuna$_i$-lul nwukwunka-ka t_i chodayhatta. ambiguous: S>O; O>S
everyone-ACC someone-NOM invited
'Someone invited everyone. (*scrambled*)'

The Korean test sentences for the present research project used *motun N* 'all the N' as the object QP, as shown in Table 41:

Table 41: Test types for Korean version of QP-QP experiment

type	sentence structure	scope
K1a:	*nwukwunka* 'someone'-NOM *motun N* 'all the N'-ACC V.PAST	S>O
K1b	*nwukwunka* 'someone'-NOM *motun N* 'all the N'-ACC V.PAST	O>S
K1c	*motun N* 'all the N'-ACC *nwukwunka* 'someone'-NOM V.PAST	S>O
K1d	*motun N* 'all the N'-ACC *nwukwunka* 'someone'-NOM V.PAST	O>S
K2a:	NumP-NOM *motun N* 'all the N'-ACC V.PAST	S>O
K2b	NumP-NOM *motun N* 'all the N'-ACC V.PAST	O>S
K2c	*motun N* 'all the N'-ACC NumP-NOM V.PAST	S>O
K2d	*motun N* 'all the N'-ACC NumP-NOM V.PAST	O>S

Five tokens of each type were judged by 22 native speakers of Korean. Their group mean ratings indicated acceptance of subject-wide scope on both SOV and OSV sentences (Types K1a and K2a, group mean ≥2.18 on the scale of 0–3; Types K1c and K2c, group mean ≥2.02). Object-wide scope was rejected on the SOV and OSV sentences (Types K1b and K2b, group mean ≤0.98; Type K1d and K2d ≤1.24). These results are thus as claimed in the theoretical literature, except for the rejection of object-wide scope on the scrambled 'd' types.

As suggested in Chapter 5, the rejection of object-wide scope on the scrambled sentences could be due to the lexical properties of *motun* 'all'. This quantifier tends to have a collective rather than a distributive interpretation (Kook-Hee Gill, personal communication, March 2004). Distributivity is required for an object-wide scope interpretation; therefore, lacking distributivity entails lacking the capacity to take object-wide scope. Consequently, the native Korean participants' rejection of object-wide scope on the scrambled 'd' sentences with *motun N* 'all the N' as object QP does not necessarily contradict the claims in the theoretical literature that object-wide scope is available in scrambled QP-QP sentences, since these claims are made about sentences with *nwukwuna* 'everyone' as object QP.

The difference between the present results for *motun N* 'all the N' and the theoretical claims about *nwukwuna* 'everyone' with respect to object-wide scope in scrambled QP-QP sentences provides further evidence of the different behaviours of different universal quantifiers. The Korean results thus highlight the need for a theory of scope interpretation that accounts for different types of quantifier. such as the Target Landing Sites theory by Beghelli (1995, 1997) and Beghelli & Stowell (1997).

6.1.1.3. Native English and native Chinese data

The native English and native Chinese data provided quantitative support for the two key claims summarised in (226) and (227) (see Chapter 2 for details):

226. English doubly-quantified SVO sentences with an existential subject and *every N* as object are scopally ambiguous, allowing both subject-wide and object-wide interpretations; however, with *all the N* as object they are unambiguous, allowing only the subject-wide interpretation.

227. Chinese doubly-quantified SVO sentences with an existential subject are unambiguous, allowing only a subject-wide interpretation, whatever the object quantifier.

The present section briefly compares the present data with those of the existing previous studies of QP-QP scope interpretation in English (Kurtzman & MacDonald 1993) and in Chinese and English (Lee *et al.* 1999b). Implications of the results for theories of scope interpretation are briefly discussed at the end of the section.

The test types used in the native English and native Chinese tasks in the present study were as shown in Table 42 (previously presented in Chapter 5; the English task used the sentences given as translations of the Chinese).

Table 42: Test types for Chinese and English versions of QP-QP experiment

type	sentence structure	scope
1a	*mouren* 'someone' *mei-CL N* 'every N'-ACC V.PAST	S>O
1b	*mouren* 'someone' *mei-CL N* 'every N'-ACC V.PAST	O>S
2a	*mouren* 'someone' *suoyoude N* 'all the N'-ACC V.PAST	S>O
2b	*mouren* 'someone' *suoyoude N* 'all the N'-ACC V.PAST	O>S
3a	NumP-NOM *mei-CL N* 'every N'-ACC V.PAST	S>O
3b	NumP-NOM *mei-CL N* 'every N'-ACC V.PAST	O>S
4a	NumP-NOM *suoyoude N* 'all the N'-ACC V.PAST	S>O
4b	NumP-NOM *suoyoude N* 'all the N'-ACC V.PAST	O>S

Five tokens of each type were judged by 24 native speakers of English and 20 native speakers of Chinese. As reported in Chapter 5, subject-wide scope was judged acceptable in both English and Chinese on all four of the 'a' types (group mean rating ≥2.68 for English, ≥2.22 for Chinese, on the scale of 0–3); in English, object-wide scope was acceptable when the object quantifier was *every* (Types 1b and 3b: group mean rating ≥1.74) but unacceptable when the object quantifier was *all* (Types 2b and 4b: group mean rating ≤0.93); in Chinese, object-wide scope was unacceptable on all four of the 'b' types (group mean rating ≤0.45).

Kurtzman & MacDonald's (1993) study of scope interpretation in English included sentences such as (228), with *a* as the subject quantifier, *every* as the object quantifier, and an active transitive verb.[193]

228. A kid climbed every tree.

These sentences are comparable with Types 3a–b in the present study (although the existential subject quantifier in 3a–b.is numerical instead of the determiner *a*). In Kurtzman & MacDonald's experiment, participants read each test sentence (such as (228)) on a computer screen, then pressed the space bar to reveal a continuation sentence. The continuation sentences required either a subject-wide interpretation of the test sentence (e.g., (229)), or an object-wide scope interpretation (e.g., (230)).

229. The kid was full of energy.

230. The kids were full of energy.

The participants pressed a 'yes' or 'no' button to indicate whether they found that the continuation 'made sense' and was a 'natural continuation' of the first sentence.

Kurtzman & MacDonald's results showed that, for sentences like (228) with an *a*-phrase subject and *every N* as object, the subject-wide continuation was accepted (by pressing the 'yes' button) far more frequently than the object-wide continuation: >80% 'yes' responses v. <25% 'yes' responses.[194] The high rate of acceptance of subject-wide scope is similar to the findings for subject-wide scope in the present study. However, the rate of <25% acceptance of object-wide scope is much lower than in the present study.[195] Kurtzman & MacDonald suggest that one factor contributing to this result, despite the theoretical ambiguity of sentences like (228), could be the *a*-phrase subject setting up a topic that refers to a single entity. In this context, the object-wide continuation, with its plural subject, is judged unnatural. The experiment design of the present study aimed to avoid such formation of preconceptions about the interpretation of the test sentences, by presenting the picture context for each sentence before presenting the actual sentence. In other

[193] Detailed presentations of Kurzman & MacDonald (1993) and Lee *et al.* (1999b) were given in Chapter 3. Only those details essential for the purpose of comparison are repeated here.
[194] Approximate percentages are taken from a chart (Kurtzman & MacDonald 1993: 255). Kurtzman & MacDonald do not provide exact values.
[195] Actual results for the present study are repeated in Table 43, following.

words, the participants saw the unambiguous picture context before they read the ambiguous sentence. Since acceptance of object-wide scope in native English was higher in the present study, this feature of the experiment design may have successfully helped to facilitate object-wide scope interpretation.

Lee *et al.*'s (1999b) investigation of QP-QP scope interpretation in English and Chinese also yielded a higher rate of acceptance of object-wide scope on English sentences with *every* as the object quantifier.[196] Among the sentence types investigated by Lee *et al.* were exemplars directly comparable with Types 3a–b (subject = NumP, object = *mei-CL N/every N,* verb = transitive) and 4a–b (subject = NumP, object = *suoyoude N/all the N,* verb = transitive) in the present study. In Lee *et al.*'s experiment, participants read each sentence along with a number of interpretations for the sentence, and rated how possible each interpretation was on a scale of 0 to 3 ('0' = 'impossible'; '3' = 'possible and often understood that way'). An example is given in (231) (Lee *et al.* 1999b: 181 (34), previously presented in Chapter 3):[197]

231. Two flags waved from every window in the building.
   A) There were only two flags. One flag waved from all the windows. The other flag also waved from all the windows. *(S>O)*
   B) There were only two flags. One flag waved from some windows. The other flag also waved from the other windows. *(cumulative interpretation)*
   C) Each of the windows had two flags waving. Different windows had different flags. There were more than two flags. *(O>S)*

The study included four tokens each of test types corresponding to Types 3a, 3b, 4a and 4b (see Table 42) in the present study. Thirty native speakers of Mandarin Chinese judged the sentences in Chinese, and 27 native speakers of English judged them in English.

Lee *et al.*'s rates of consistent acceptance of object-wide scope for each participant group are presented in Table 43, and compared with the rates of consistent acceptance, consistent rejection, and inconsistency for all four object-wide scope types (Types 1b, 2b, 3b and 4b) in the present study. (Since there were four tokens per type in Lee *et al.* (1999) and five per type in the present study, the consistency criterion is slightly stricter in the present study. See the Table 5 notes.)

[196] Actual results are repeated in Table 43, following.
[197] The example in (231) has an unaccusative verb because Lee *et al.* do not provide an example of their test types with active transitive verbs.

Table 43: Comparison of native Chinese and native English consistency on object-wide scope types in the present study and in Lee *et al.* (1999b)

type	subject QP; object QP (scope = O>S) (verb = transitive)	Present study						Lee *et al.* (1999b)	
		Chinese (n=20)			English (n=24)			Chinese (n=30)	English (n=27)
		acc[a] % (n)	rej[b] % (n)	inc[c] % (n)	acc[a] % (n)	rej[b] % (n)	inc[c] % (n)	acc[d] %	acc[d] %
1b	*mouren/someone*; *mei-CL N/every N*	0.0 (0)	100.0 (20)	0.0 (0)	54.2 (13)	12.5 (3)	33.3 (8)	—	—
2b	*mouren/someone*; *suoyoude N/all the N*	5.0 (1)	95.0 (19)	0.0 (0)	12.5 (3)	70.8 (17)	16.7 (4)	—	—
3b	NumP; *mei-CL N/every N*	5.0 (1)	90.0 (18)	5.0 (1)	62.5 (15)	20.8 (5)	16.7 (4)	26.7	44.4
4b	NumP; *suoyoude N/all the N*	0.0 (0)	90.0 (18)	10.0 (2)	8.3 (2)	75.0 (18)	16.7 (4)	10.0	11.1

[a] 'acc' = 'consistent acceptance': % (n) of participants selecting ratings of 2 (+1) or 3 (+2) on at least four of the five test tokens.
[b] 'rej' = 'consistent rejection': % (n) of participants selecting ratings of 0 (–2) or 1 (–1) on at least four of the five test tokens.
[c] 'inc' = 'inconsistency': % (n) of participants with ratings corresponding neither to consistent acceptance, nor to consistent rejection.
[d] 'acc' = 'consistent acceptance': % of participants selecting ratings of 2 or 3 on at least three of the four test tokens.

Table 43 shows that Lee *et al.*'s consistent acceptance ratings on Type 3b (subject = NumP, object = *mei-CL N/every N*) and Type 4b (subject = NumP, object = *suoyoude N/all the N*) are broadly similar to those from the present study in that the highest rating in each case occurs in the native English data when the object quantifier is *every* (Type 3b): 62.5% in the present study; 44.4% in Lee *et al.*'s study. When the object quantifier is *all* (Type 4b), consistent acceptance ratings are much lower: 8.3% in the present study; 11.1% in Lee *et al.*'s study. The native Chinese consistent acceptance ratings on both Types 3b (object = *mei-CL N* 'every N')) and 4b (object = *suoyoude N* 'all the N') are also lower in both studies than the native English rates for Type 3b. However, the rates in the present study are extremely low: 5% on Type 3b, 0% on Type 4b; in Lee *et al.*'s study, the corresponding native Chinese rate for Type 3b is considerably higher at 26.7% , and the Type 4b rate is also somewhat higher, at 10.0%. The consistent acceptance ratings for Type 1b (subject = *mouren/someone*, object = *mei-CL N/every N*) and Type 2b (subject = *mouren/someone*, object = *suoyoude N/all the N*) in the present study are similar to those for Types 3b and 4b: native Chinese acceptance ratings are extremely low (0% on Type 1b; 5% on Type 2b); native English acceptance is relatively high when the object quantifier is *every* (Type 1b: 54.2%) but much lower when the object quantifier is *all* (Type 2b: 12.5%).

Overall, the results from the present study corroborate Lee *et al.*'s findings, while, at the same time, testifying more clearly than theirs to the claims stated in (226), *viz.*, that English allows object-wide scope with *every N* but not with *all the N*, and (227), *viz.*, that object-wide scope is not readily available in Chinese. The reason for the clearer results in the present study is probably due to the difference in experimental designs. It was observed in Chapter 3 that Lee *et al.*'s design was very reading-intensive, and also very long (100 interpretations were judged by each participant). Fatigue, and confusion due to puzzling over 100 written scope interpretations, may have impaired participants' judgements. One of the aims of the design adopted in the present study was to minimise the test-taking burden in order to maximise the accuracy of the judgements. The clearer results of the present study compared with Lee *et al.*'s suggest that this aim was successfully realised.

In spite of these efforts to facilitate clear judgements about scope interpretation, the native English data from the present study still include evidence of a degree of inconsistency, as with the native Japanese results (although the native Chinese data from the present study are almost inconsistency-free). Specifically, the native English speakers did not unanimously accept object-wide scope when the object quantifier was *every* (mean consistent acceptance on Type 1b = 54.2%, Type 3b = 62.5%); nor did they unanimously reject object-wide scope when the object quantifier was *all* (mean consistent rejection on Type 2b = 70.8%, Type 4b = 75%).

The lack of consistent acceptance of object-wide scope when the object quantifier is *every* is likely to be due to the fact that object-wide scope interpretations are harder to process than subject-wide scope interpretations, even when genuine ambiguity exists (as Kurtzman & MacDonald's (1993) study testifies; see details in Chapter 3). Such a lack is not problematic for the theoretical accounts of ambiguity in English QP-QP sentences, given the overall pattern in the data of acceptance of object-wide scope when the object quantifier is *every*.

The results for English QP-QP sentences with *all* as the object quantifier are more problematic for the accounts of quantifier scope detailed in Chapter 2. First, before considering the inconsistencies in the results for *all*, the overall result showing rejection of object-wide scope when the object quantifier is *all* provides further quantitative evidence of the need for accounts of quantifier scope to address the differences between different quantifiers: the theories of May (1977, 1985), Aoun & Li (1993) and Hornstein (1995) focus predominantly on the quantifiers *some* and *every*, with an implication that other existential and universal quantifiers exhibit the same patterns. These theories thus do not account for the lack of object-wide scope of *all*. The Target Landing Sites theory by Beghelli (1995, 1997) and Beghelli & Stowell (1997), on the other hand, specifically addresses differences between quantifiers. Under this theory, it is predicted that object-wide scope is not possible for *all* in the sentence types investigated. The overall result thus supports this theory. However, the consistent acceptance of object-wide scope for *all* by a minority of participants (3 out of 24 on Type 2b, subject = *someone*; two out of 24 on Type 4b, subject = NumP) is not predicted. Modification of a constrained theory, such as the Target Landing Sites theory, in order to account for a small minority of unpredicted experimental results is clearly not desirable, since such a modification would decrease the explanatory power of the theory. As with the four instances of consistent acceptance of object-wide scope in the native Japanese data, it seems that the instances of acceptance of object-wide scope for *all* in the English data must be written off as a quantitative expression of the overall complexity of scope interpretations.

6.1.2. *Wh*-QP study: native experimental data

The *Wh*-QP study investigated questions with a QP subject and *wh*-object in Japanese, Korean, Chinese, English, as exemplified in (232)–(234) (the English translations exemplify the question type used in the English version of the task):[198]

232. *Japanese*:
Nani-o daremo-ga katta no?
what-ACC everyone-NOM buy.PAST Q
'What did everyone buy? (*scrambled*)'

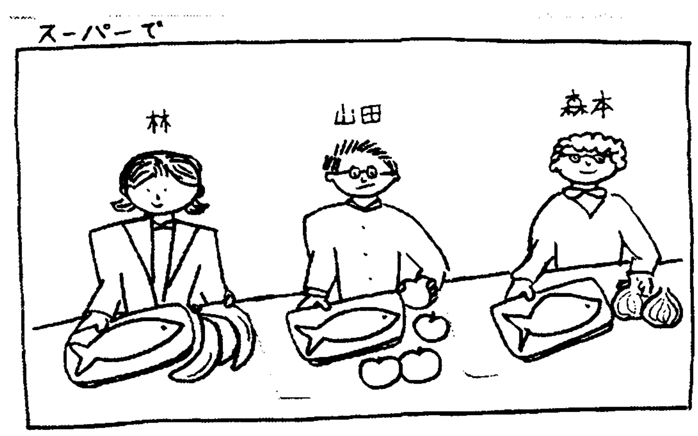
233. *Korean:*
Mwues-ul nwukwuna-ka sass ni?
what-ACC everyone-NOM buy.PAST Q
'What did everyone buy? (*scrambled*)'

[198] Scrambled questions were used in Japanese and Korean because of the reported marginal acceptability of the non-scrambled versions of *wh*-QP questions with a *wh*+QPt subject (Hoji 1985; Kim 1989; Tomioka 2004). See Chapter 2 for details.

234. *Chinese*:
Meigeren maile shenme?
everyone bought what
'What did everyone buy?

Participants made judgements about the acceptability of individual answers (235a) and pair-list answers (235b) to questions of the types illustrated in (232)–(234). As shown in (235), a picture provided a plausible context for both the individual interpretation and the pair-list interpretation. (The example picture in (235) includes text in Japanese, and the example answers in (235a–b) are provided in English. However, the actual test materials were fully 'localised' for each language. In addition, the English translation of the text in the picture, provided for convenience in (235), was not included in the actual test materials. See Chapter 5 for full details of the procedure, and Appendix 6 for the test materials.)

235. *Picture context:*



Translation (top): At the supermarket
(above heads): Hayashi; Yamada; Morimoto

a. *Individual answer:*
Fish.

b. *Pair-list answer:*
Mrs Hayashi bought fish and bananas, Mr Yamada bought fish and apples, and Mrs Morimoto bought fish and onions.

Five tokens of each of the answer types in (235) were judged by 18 native speakers of Japanese, 25 native speakers of Korean, 14 native speakers of Chinese, and 18 native speakers of English, all in their respective native language.

The results indicated that individual answers were acceptable in native Japanese, native Korean and native English (group mean ratings ≥2.01 on scale of 0–3) but unacceptable in native Chinese (group mean: 1.27); and pair-list answers were acceptable in native Korean, Chinese and English (group mean ratings ≥2.1) but unacceptable in native Japanese (group mean: 1.29). As noted in Chapter 5, two of these results contradict the claims of the theoretical literature (see Chapter 2 for details): individual answers are claimed to be available in native Chinese (e.g., Aoun & Li 1993); and pair-list answers are claimed to be unavailable in native Korean (Beck & Kim 1997; Kim 1989).[199]

The unexpected rejection of individual answers in Chinese was discussed in Chapter 5.[200] The source of the unexpected acceptance of pair-list answers in Korean is explored in Section 6.2.2.2. Before that, Section 6.2.2.1 briefly discusses the implications of the individual consistency patterns in the native Japanese results.

6.1.2.1. Individual consistency in native Japanese

It was shown in Chapter 5 that, although pair-list answers were rejected in group terms in native Japanese, 6 of the 18 native Japanese participants demonstrated consistent acceptance of pair-list answers. The individual consistency patterns for the native Japanese group are repeated in Table 44, and compared with those of the other three groups.

---

[199] Note that Beck & Kim (1997) and Kim (1989) refer only briefly to a lack of pair-list interpretations in scrambled [*wh*-object...*nwukwuna-ka* 'everyone-NOM'...] questions in Korean. To the author's knowledge, there is no in-depth exploration of these questions in the theoretical literature on Korean.

[200] Briefly, Aoun & Li's (1993) claim that Chinese allows both individual and pair-list answers is made about questions that include the floating quantifier *dou* 'all', such as (i).

i. Meigeren dou maile shenme?
   everyone all bought what
   'What did everyone buy?'

The Chinese question form (234) used in the present study did not include *dou*. See Chapter 5 for further discussion.

Table 44: Individual consistency on *wh*-QP task in native Japanese, English, Chinese and Korean

group	individual answers			pair-list answers		
	acc[a] % (n)	rej[b] % (n)	inc[c] % (n)	acc[a] % (n)	rej[b] % (n)	inc[c] % (n)
JJ (n=18)	94.4 (17)	5.6 (1)	0	33.3 (6)	55.6 (10)	11.1 (2)
EE (n=21)	81.0 (17)	0	19.0 (4)	100 (21)	0	0
CC (n=14)	14.3 (2)	50.0 (7)	35.7 (5)	85.7 (12)	0	14.3 (2)
KK (n=25)	76.0 (19)	8.0 (2)	16.0 (4)	72.0 (18)	4.0 (1)	24.0 (6)

[a] 'acc' = 'consistent acceptance': % (n) of participants selecting ratings of 2 (+1) or 3 (+2) on at least four of the five test tokens.
[b] 'rej' = 'consistent rejection': % (n) of participants selecting ratings of 0 (–2) or 1 (–1) on at least four of the five test tokens.
[c] 'inc' = 'inconsistency': % (n) of participants demonstrating neither consistent acceptance nor consistent rejection.

The only pattern that does not demonstrate at least 72% consistent acceptance or rejection of the particular answer type, apart from the native Chinese group's pattern for individual answers, is the native Japanese response pattern for pair-list answers.[201] Although the majority of native Japanese participants consistently rejected pair-list answers, this majority is only 10 out of 18 participants (55.6%). Of the remaining eight participants in the group, six (33.3%) consistently accepted pair-list answers and two (11.1%) had inconsistent response patterns.

This set of response patterns seems to testify to (i) pair-list answers being generally unacceptable in Japanese; but (ii) some variety (or varieties) of Japanese allowing pair-list readings. The native Japanese result contrasts so starkly with the results for pair-list answers by the other three groups (where pair-list answers are consistently accepted by at least 72% of the group) that it seems reasonable to maintain a general categorisation of Japanese as disallowing pair-list readings while the other three languages allow them. However, given that six of the native Japanese participants consistently allowed pair-list readings, some processing mechanism must allow the general absence of pair-list readings to be over-ruled. Further

[201] The inconsistency by the native Chinese participants is likely to be due to the fact that the question form used in the Chinese version of the task (see 234) sounded somewhat unnatural in native Chinese. This was discussed in Chapter 5 and is not pursued further here.

exploration of the processing of scrambled [*wh*-object...QP-subject...] questions in Japanese is left for future research.

6.1.2.2. Native Korean data

One possible source of the unexpected acceptance of pair-list interpretations in scrambled Korean [*wh*-object...QP-subject...] questions could be the properties of the Korean quantifier *nwukwuna* 'everyone', used as the QP-subject. It was expected that Korean would be like Japanese, with pair-list interpretations being rejected. However, the Korean *wh*+QPt quantifier *nwukwuna* is not directly equivalent to the quantifier used as the QP-subject in the Japanese questions, *daremo*, even though they are frequently presented as counterparts in questions with 'everyone' as the subject and a *wh*-word as the object (e.g., Beck & Kim 1997; Kim 1989; Tomioka 2004). The extensive investigation of *wh*+QPt quantifiers in Japanese and Korean by Gill (2002) shows that Korean *nwukwuna* and Japanese *daremo* are compositionally and distributionally different. Two further quantifiers are introduced—*daredemo* 'anyone/everyone' in Japanese, and *nwukwuto* 'no-one' in Korean—in order to fully illustrate the difference.

Considering Japanese *daremo*, first, compositionally, this is the equivalent of Korean *nwukwuto* 'no-one'—and not *nwukwuna* 'everyone'—since each has the structure given in (236):[202]

236. 'who' (*dare/nwukwu*) + conjunctive particle (*mo/to*)

However, distributionally, *daremo* and *nwukwuto* are not exactly the same. Notably, *nwukwuto* is licensed only by a negative environment, as illustrated in (237) and (238), where (237) shows the grammatical usage with a negated verb, and (238) is ungrammatical due to the affirmative context (based on Gill 2002: 12 (16–17)):

237. Nwukwuto tapcang-ul acik anh hayssta.
nobody reply-ACC yet NEG did
'Nobody replied yet.'

[202] Examples of the conjunctive use of the particles *mo* (in Japanese) and *to* (in Korean) are as follows:

i. *Japanese*: Sam-mo Sally-mo
Korean: Sam-to Sally-to
'Both Sam and Sally'

238. *Nwukwuto pap-ul masisskey mantulessta.
nobody rice-ACC deliciously made

By contrast, Japanese *daremo* has two usages: it is negative-sensitive in the same way as Korean *nwukwuto* when it is used without a case marker, as shown in (239); however, when used with a case marker—as seen throughout this dissertation—it is not polarity-sensitive, as shown in (240) (Hasegawa 1993; Nishigauchi 1990, 1999).

239. Daremo ko-na-katta/*kita.
nobody come-NEG-PAST/came
'Nobody came./*Anybody came'

240. Daremo-ga ko-na-katta/kita.
everyone-NOM come-NEG-PAST/came
'Everyone didn't come./Everyone came'

Gill (2002: 18) states that, for Korean *nwukwuto,* non-polarity-sensitive usage with a case marker (analogous with Japanese *daremo* in (240)) is not possible. Therefore, *nwukwuto* is ungrammatical as the subject of a scrambled [*wh*-object...QP-subject...] question such as those investigated in this dissertation, as shown in (241).

241. *Mwues-ul nwukwuto(-ka) sass ni?
what-ACC no-one(-NOM) bought Q

Moving on to Korean *nwukwuna,* this differs compositionally from *daremo/nwukwuto* in two ways. First, the Korean quantificational particle *(i)na* is a disjunction morpheme, not a conjunction morpheme.[203] Second, the structure of *nwukwuna* is claimed to incorporate the Korean copula. Specifically, the *–i* of *(i)na,* which surfaces post-consonantally, is argued to be a form of the Korean copula (Jang 1999; Lee 1996; Martin 1992). Thus, the structure of *nwukwuna* is as shown in (242a), while (242b) provides an example of a Korean *wh*+QPt quantifier in which the *–i* is overt. As indicated by the translations in (242), Korean *wh+(i)na* quantifiers have the 'free choice' sense of English *any*, as well as the universal sense of English *every*, although this dissertation has focused thus far only on the universal sense.

242. a, nwukwu-(i)-na
who-(COP)-DISJ
'anyone/everyone'

[203] For example, *Chelswu-na Younghi* in Korean means 'Chelswu or Younghi'.

b. mwues-i-na
what-COP-DISJ
'anything/everything'

The Japanese quantifier *daredemo* 'anyone/everyone'—and not *daremo*—is like *nwukwuna* in that it can be analysed as incorporating a copula morpheme, *de*. However, *daredemo* differs from *nwukwuna* in that the quantificational particle is the conjunction *mo*, not a disjunction, like Korean *na*. The composition of *daredemo* is shown in (243).

243. dare-de-mo
who-COP-CONJ

Neither Korean *nwukwuna* nor Japanese *daredemo* are polarity sensitive. Regarding meaning, Gill (2002) reports, based on a database of native-speaker judgements, that both can receive free-choice, universal, and existential interpretations. However, it seems that in both languages, the free-choice meaning, corresponding (in broad terms) to English *anyone* is the most characteristic usage. For example, Nishigauchi (1999: 281) defines Japanese *daredemo* as a 'free-choice indefinite'; and Gill, Harlow & Tsoulas (2003: 13 (37)) report that native Korean informants usually exemplify the meaning of *nwukwuna* with a sentence in which the free-choice reading is the most salient, such as (244):

244. Nwukwuna ke kes-ul halswuissta.
anyone the thing-acc can-do
'Anyone can do it.'

Moreover, Japanese *daredemo* differs from Korean *nwukwuna* in that it is ungrammatical in scrambled [*wh*-object...QP-subject...] questions:[204]

245. *Nani-o daredemo-ga katta no?
what-ACC anyone-NOM bought Q

In other words, despite the similarities between Korean *nwukwuna* and Japanese *daredemo*, they differ distributionally on precisely the question type investigated in

[204] Thanks to Hidekazu Tanaka for useful discussion (June 2004) of the grammaticality of *daredemo*. Note that (245) is not improved by omitting the case marker *–ga*, even though *daredemo* can occur without a case marker.

this dissertation: *nwukwuna*, but not *daredemo*, is grammatical as the subject in a scrambled [*wh*-object... 'everyone'-subject...] question.

The morphology of the four quantifiers is summarised in Table 45. along with those details of the meaning and distribution that are pertinent to the discussion.

Table 45: Summary of morphology, meaning and distribution of four quantifiers in Japanese and Korean

quantifier	morphology	meaning & distribution
Japanese: dare-mo	'who'+CONJ	'no-one': negative environment, no case-marking; 'everyone': case-marking required; *wh*-ACC ...-NOM V?
Korean: nwukwu-to	'who'+ CONJ	'no-one': negative environment, no case-marking; *...-CASE; **wh*-ACC ...(-NOM) V?
Japanese: dare-de-mo	'who'+ COP + CONJ	'anyone'/(everyone'); *wh-ACC ...(-NOM) V?
Korean: nwukwu-(i)-na	'who'+ (COP) + DISJ	'anyone/everyone'; wh-ACC ...(-NOM) V?

Table 45 clarifies that none of the four quantifiers is directly equivalent to any other. Most importantly for the present discussion, Japanese *daremo* is far from being directly equivalent to Korean *nwukwuna*, despite the fact that the two have been presented in the literature as equivalents.

The reason why *daremo* and *nwukwuna* have nonetheless been presented as equivalents in scrambled [*wh*-object...QP-subject...] questions is, no doubt, that, of the four *wh*+QPt quantifiers in Table 45, *daremo* and *nwukwuna* are grammatical, with the meaning of 'everyone', in this context.[205] However, the subtle differences between *daremo* and *nwukwuna* may be the source of the finding in the present study that pair-list interpretations were judged acceptable in Korean [*wh*-object... *nwukwuna*-NOM...] questions, but they were judged unacceptable in Japanese [*wh*-object... *daremo*-NOM...]. Specifically, the free-choice meaning of *nwukwuna* may play a role. The example Korean test question given in (233) at the beginning of this section is repeated in (246), but this time *nwukwuna* is translated as 'anyone' instead of 'everyone'.

[205] As pointed out elsewhere, both Korean and Japanese also have non-*wh*+QPt quantifiers with the meaning of 'everyone'. Suh (1990) discusses scrambled Korean [*wh*-object...QP-subject...] questions with *motun saram* 'all the people' as the subject QP, and states that pair-list readings are unavailable. This is unsurprising if, as discussed above (Section 6.1.1.2), the interpretation of Korean *motun* 'all' is predominantly collective rather than distributive. Further examples of non-*wh*-QPt universals are *motwu* in Korean and *minna* in Japanese. These may also lack the distributive sense of the *wh*+QPt universals. They are not discussed here.

246. Mwues-ul nwukwuna-ka sass ni?
what-ACC anyone-NOM buy.PAST Q
'What did anyone buy? (*scrambled*)'

The question *What did anyone buy?* is hard to interpret in English.[206] However, if an interpretation is forced, a pair-list answer, in which each individual is mapped to the thing she/he bought, is certainly possible, as exemplified in (247).

247. *Context:* Person X is holding a garage sale. She leaves the sale for an hour and Person Y takes charge of it. The following Q/A takes place on X's return.
X: What did anyone buy?
Y: Someone bought that old clock, someone bought a few books, the lady from down the road bought the armchair, etc... .

The example in (247) is intended only to illustrate that a pair-list interpretation of a question with a *wh*-object and 'anyone' as the subject is conceptually possible. Further investigation is required in order to discover whether interpretation of *nwukwuna* as 'anyone' instead of 'everyone' is the reason why the native Korean informants accepted pair-list interpretations of questions like (246) in the judgement task. The fact that Korean *nwukwuna* allows a free-choice interpretation as well as a universal interpretation whereas Japanese *daremo* is (predominantly) universal, seems like a promising avenue to explore for the source of the unexpected acceptance of the pair-list reading in Korean and its rejection in Japanese.

6.2. Implications of the L2 results for theories of quantifier scope interpretation

6.2.1. QP-QP scope interpretation

The rationale for the investigation of QP-QP scope interpretation in L2 Japanese rests upon the fact that English allows an object-wide scope interpretation of [S...O...] QP-QP sentences while Japanese does not. Acquisition of the lack of object-wide scope in Japanese SOV QP-QP sentences is a poverty-of-the-stimulus problem for native English-speaking learners (as set out in detail in Chapter 2). As such, the only way in which English-speaking learners can acquire native-like knowledge of the lack of object-wide scope in Japanese is by indirect evidence in the

[206] The same is likely to be true in Korean: 5 of the 25 native Korean test participants commented on their answer sheet that the questions in the *wh*-QP task were 'odd' or 'difficult to understand'.

input triggering UG-constrained restructuring of the English-based interlanguage grammar. What exactly the relevant trigger may be depends on how the difference in scope interpretation between Japanese and English is accounted for. Chapter 2 included an assessment of what might trigger the relevant English-Japanese interlanguage restructuring under five different theories of quantifier scope interpretation: Hoji (1985); Aoun & Li (1993); Hornstein (1995); the Target Landing Sites theory by Beghelli (1995, 1997) and Beghelli & Stowell (1997); and Hayashishita's (1999), (2000a), (2000b) theory enhanced by the proposal by Tomioka (2004) that Japanese *wh*+QPt quantifiers are incompatible with topic-hood (henceforth, in this section, 'Hayashishita/Tomioka'). In addition, a prediction was made about the L2 learnability of Japanese scope rigidity by English-speaking learners under each account. To briefly restate the predictions (see Chapter 2 for details), it was argued that, for the first three accounts (Hoji 1985; Aoun & Li 1993; Hornstein 1995), English-speaking learners would be unable to overcome poverty of the stimulus: they would fail to acquire the lack of object-wide scope in Japanese SOV QP-QP sentences. However, under the remaining two accounts (the Target Landing Sites theory and Hayashishita/Tomioka) the relevant acquisition would be possible, but probably slow.

'Slow, but possible' acquisition is indeed what the experimental results revealed. Intermediate English-speaking learners did not appear to have knowledge of the lack of object object-wide scope in Japanese SOV QP-QP sentences: they allowed non-native-like inverse scope interpretations. On the other hand, advanced English-speaking learners behaved more like native speakers of Japanese. Specifically, six of the nine advanced English-speaking learners consistently rejected object-wide scope in Japanese SOV QP-QP sentences. This result is thus compatible with the L2 learnability predictions under the Target Landing Sites theory and Hayashishita/Tomioka, but not with the prediction, under Hoji (1985), Aoun & Li (1993) and Hornstein (1995), that the poverty-of-the-stimulus problem cannot be overcome.

Though compatible with the Target Landing Sites theory and Hayashishita/Tomioka, the L2 results do not shed any light on which (if either) of the two is more likely to be correct.[207] In Chapter 2, it was suggested that, under the

[207] The experiments were not designed to differentiate between the different theoretical accounts of scope interpretation, since this was not a goal of the present study.

Target Landing Sites theory, acquisition by native English-speakers of the lack of object-wide scope in Japanese SOV QP-QP sentences should be a corollary of acquiring the knowledge that Japanese lacks a morpho-syntactic singular/plural distinction; while within the Hayashishita/Tomioka framework, the absence of object-wide scope may be linked with the inability of Japanese *wh*+QPt quantifiers to act as discourse topics. A useful experiment that could potentially differentiate between the two accounts would be an investigation of the order of acquisition in English-Japanese interlanguage of (i) the lack of object-wide scope in Japanese and (iia) knowledge of the Japanese lack of singular/plural distinction or (iib) knowledge of the incompatibility of *wh*+QPt quantifiers with topic-hood. If learners gain native-like knowledge of Japanese scope rigidity before they gain native-like knowledge of the lack of singular/plural distinction in Japanese, this would detract from the Target Landing Sites theory. On the other hand, acquisition of native-like knowledge of Japanese scope rigidity before acquisition of native-like knowledge of the incompatibility of *wh*+QPt quantifiers with topic-hood would detract from the Hayashishita/Tomioka account.

6.2.2. *Wh*-QP scope interpretation

With regard to the interpretation of [*wh*-object...QP-subject...] questions, only the analysis by Aoun & Li (1993) appeared to account for the availability of pair-list interpretations in English and Chinese, and the lack of such interpretations in Japanese.[208] Under this account, the lack of pair-list interpretation in Japanese was linked to the fact that Japanese makes use of a sentence-final question-marking particle (predominantly *ka* or *no*) in question formation (see Chapter 2 for details). Assuming this to be correct, the L2 learnability problem for L1 English- and L1 Chinese-speaking learners of Japanese was stated in terms of coming to know that Japanese is a 'question particle language', in contrast to Chinese and English, which do not use question particles. It was suggested that question particles in the Japanese input would serve as clear and readily-available evidence that Japanese is a question particle language, and would trigger the relevant restructuring. Thus, acquisition of the lack of pair-list interpretations of scrambled Japanese [*wh*-object...QP-subject...]

[208] The discussion here focuses initially on the implications of the findings about *wh*-QP interpretation in English-Japanese and Chinese-Japanese interlanguage. The native Korean and Korean-Japanese interlanguage findings are broached at end of this section.

questions was predicted to be possible in English-Japanese and Chinese-Japanese interlanguage, despite poverty of the stimulus.

The results presented in Chapter 5 indicated that some advanced L1 Chinese- and L1 English-speaking learners did indeed have knowledge of the lack of pair-list interpretations in Japanese. The numbers of individuals who consistently rejected pair-list interpretations (i.e., who demonstrated native-like behaviour) were small: four out of 10 advanced Chinese-speaking learners of Japanese (40%) and four out of nine advanced English-speaking learners of Japanese (44.4%). Nonetheless, the proportion of learners in each advanced group who consistently rejected pair-list interpretations was larger than in each intermediate group: two out of 19 intermediate English-speaking learners (10.5%); one out of six intermediate Chinese-speaking learners (14.3%). This difference between the advanced and the intermediate groups was interpreted to show that knowledge of the lack of pair-list interpretations in scrambled Japanese [*wh*-object...QP-subject...] questions can arise despite poverty of the stimulus. However, the fact that fewer than half of the advanced English-speaking and Chinese-speaking learners acquired this knowledge indicates that the evidence to motivate the relevant restructuring must be very sparse or very obscure. This suggests that Japanese question markers in the input do not serve as sufficient evidence to motivate restructuring of the English-Japanese and Chinese-Japanese interlanguages to rule out pair-list interpretations. If evidence from question markers alone were sufficient to motivate this restructuring, knowledge of the lack of pair-list interpretations in Japanese would be expected to be demonstrated by more of the learners, since question markers are frequent and (apparently) obvious in the input.

Although it was predicted under Aoun & Li's (1993) analysis of *wh*-QP interpretation that the lack of a pair-list interpretation of scrambled Japanese [*wh*-object...QP-subject...] questions would be easily acquired by Chinese-speaking and English-speaking learners, the fact that this prediction was not confirmed (i.e., that the lack of pair-list interpretation appears hard to acquire) does not constitute evidence that Aoun & Li's account is incorrect. The continued acceptance of pair-list interpretations by more than half of the advanced learners could arise because the Japanese 'question particle grammar' can somehow be (incorrectly) represented by the English-Japanese and Chinese-Japanese interlanguage grammars without restructuring. If this is the case, Aoun & Li's analysis remains viable.

However, a more serious challenge to Aoun & Li's analysis comes from the native Korean data. As already discussed in Section 6.1.2, it was discovered that native speakers of Korean readily accept a pair-list interpretation of scrambled Korean [*wh*-object...QP-subject...] questions, despite the fact that Korean makes use of question-marking particles, like Japanese. The proposed link of languages that use question-marking particles with the unavailability of pair-list interpretations in [*wh*-object...QP-subject...] questions thus does not appear to hold.

In short, although the L2 acquisition data do not detract from Aoun & Li's (1993) analysis of the interpretation of [*wh*-object...QP-subject...] questions, the native Korean data do. Since the other theories of *wh*-QP interpretation outlined in Chapter 2 accounted only for the availability of pair-list interpretations in English and not for the lack thereof in Japanese, it is clear that more research is required to explain the complex cross-linguistic differences attested to in the experimental data.

6.3. Comments on the methodology

The present study adopted the methodological framework proposed by Schwartz & Sprouse (2000): a comparative interlanguage study incorporating L2 poverty-of-the-stimulus problems. Schwartz & Sprouse (2000) argue that experimentation within this framework should provide a clear window on the roles of L1 transfer and UG in L2 acquisition. Specifically, investigation of L2 development of a phenomenon, *P*, by learners whose L1s differ with respect to that phenomenon sheds light on the role of L1 transfer in L2 acquisition; and investigation of whether learners can acquire *P* when *P* represents a poverty-of-the-stimulus problem sheds light on the role of UG. An advantage of the framework is that the results remain valid regardless of any changes in the technicalities of the theories accounting for *P*.

The results obtained in the present study testify to the utility of this framework. 'Developmental paths' were investigated by comparing learners at an intermediate level of L2 Japanese proficiency with learners at an advanced level. In the investigation of QP-QP interpretation, the developmental path of native English-speaking learners of Japanese differed from the path of native Chinese-speaking and native Korean-speaking learners of Japanese in precisely the way that L1 transfer predicts: intermediate-level English-speaking learners allowed non-target-like object-wide scope on Japanese SOV QP-QP sentences, while intermediate-level Korean-speaking and Chinese-speaking learners did not. This was argued to reflect the facts of the L1s—object-wide scope is allowed in English QP-QP sentences, but not Chinese and Korean—and thereby to provide evidence of L1 transfer at an early stage of L2 acquisition. Similarly, in the investigation of the interpretation of questions with a QP subject and *wh*-object, the intermediate-level native Chinese-speaking learners differed from the intermediate English-speaking and Korean-speaking learners in that the Chinese group demonstrated a depressed rate of acceptance of 'individual' interpretations (i.e., QP-subject > *wh*-object interpretations). This was argued to reflect the unacceptability of individual answers in native Chinese compared with the high acceptability of individual answers in native English and Korean, and thus, again, provide evidence of L1 transfer.[209] Regarding poverty of the stimulus, in the QP-QP study, acquisition of the lack of object-wide scope in Japanese SOV QP-QP sentences was a poverty-of-the-stimulus problem for English-speaking learners; and in the *wh*-QP study, acquisition of the lack of pair-list interpretations in scrambled [*wh*-object...QP-subject...] questions was a poverty-of the-stimulus problem for all three learner groups. For both phenomena, subsets of the advanced learner groups demonstrated target-like behaviour, indicating that L2 poverty of the stimulus can be overcome. This was argued to implicate UG in the L2 acquisition process. The overall conclusion—that the L1 grammar and UG are privileged sources of knowledge in adult L2 acquisition—can be maintained whatever theory accounts for the scope interpretation phenomena investigated.

The results of the present study show that conducting research within Schwartz & Sprouse's (2000) framework facilitates clear, useful conclusions about L2 acquisition theory. If, for example, the 'comparative' aspect of the design had not been followed, and only English learners of Japanese had been investigated, the fact that the English learners' developmental path for QP-QP interpretation reflects QP-QP interpretation in English could have been argued to indicate a universal L2 acquisition pattern for scope interpretation (i.e., all L2 learners initially allow object-wide scope in their L2) without reference to L1 influence. However, since English-speaking learners were compared with Chinese-speaking and Korean-speaking learners, and the developmental paths of all three groups reflected the divergence

[209] Recall that, at the outset of the research, it was expected that native Korean would differ from native Chinese and English in that the former would not allow pair-list interpretations. However, as detailed in Chapter 5 and discussed in the present chapter, the native Korean control data did not attest to such a difference.

between the L1s, the argument that this finding is due to L1 influence, and not to some universal L2 acquisition pattern, becomes much stronger.

A serious practical consideration concerning Schwartz & Sprouse's methodology is the complexity it forces on the experiment design. It requires multiple participant groups: learner groups are required of at least two L1 backgrounds and at least two proficiency levels; in addition, native control data in both the target language and the L1s considerably enhances the study.[210] This means that, as well as the challenge of sourcing the necessary numbers of participants and of timetabling and conducting the experiments, a crucial feature of the test instrument design is that it must be suitable for translation into all the languages of the study. In addition, reliable help with the important task of translation must be sought. Ideally, pilot tests should be conducted in all of the languages under investigation. In the present study, there was no full-scale pilot of the Chinese and Korean materials due to constraints of time, finance, and access to suitable pilot test participants. Had it been possible to pilot all materials, the unexpected results in the Chinese and Korean *wh*-QP tasks (i.e., the native Chinese rejection of individual answers and the native Korean acceptance of pair-list answers) would have been discovered earlier, enabling the formulation of appropriate L2 acquisition hypotheses at the outset of the main study. Finally, the statistical analysis of the results from a study within this framework is likely to be complex, given the large number of groups that must be compared with respect to at least one variable with at least two levels.

All of the challenges enumerated here can be overcome, and they are not unique to Schwartz & Sprouse's methodology: most L2 acquisition studies entail some or all of these considerations. Overall, the methodology provides a clear framework for yielding informative L2 acquisition data. Further experimentation within this framework would be beneficial for the field of L2 acquisition research.

[210] Collection of native control data in the L1s could conceivably be dispensed with if reliable relevant quantitative data are already available. However, in the present study, the native control data proved crucial for understanding the L2 results in the *wh*-QP study: the fact that pair-list interpretations were accepted in intermediate L1 Korean-L2 Japanese interlanguage would have been very puzzling without the native Korean data showing that pair-list interpretations are acceptable in equivalent native Korean questions.

6.4. Conclusion

This chapter has discussed a number of issues arising from the main experimental studies conducted for this dissertation. Section 6.1 reviewed the native control data and highlighted the following four issues: (i) the native Japanese data on QP-QP interpretation clearly confirmed that object-wide scope is not available in Japanese SOV QP-QP sentences with an active, transitive verb. This quantitative evidence poses a challenge for the theories of Kuno & Takami (2002) and Hayashishita (1999, 2000a, 2000b), which treat Japanese QP-QP sentences as allowing object-wide scope; (ii) the native English and Chinese QP-QP data clearly showed rejection of object-wide scope in native Chinese and acceptance of object-wide scope in English when the object quantifier was *every*. This result was shown to corroborate the findings of Lee *et al.* (1999b), and also to surpass the latter study in robustness of the results. It was suggested that the reason for the clearer results in the present study was the simpler task required of the participants; (iii) a degree of inter-participant variation was identified in all of the native results, whereby a minority of participants accepted/rejected one scope reading while the rest of the group rejected/accepted it. It was suggested that this is a quantitative reflection of the subtlety and fragility of quantifier scope judgements about scope interpretation, and may be unavoidable in quantitative research on scope interpretation; (iv) pair-list interpretations of scrambled [*wh*-object...QP-subject...] questions were accepted in native Korean, contra the claims of the theoretical literature. It was suggested that this result may be due to the lexical properties of the quantifier *nwukwuna* 'anyone/everyone' (cf. Japanese *daremo* 'everyone') used as the subject QP in the Korean *wh*-QP task.

Section 6.2 discussed the implications of the L2 acquisition data for theories of scope interpretation. Under two of the theories of quantifier scope interpretation discussed in this dissertation (the Target Landing Sites theory (Beghelli 1995, 1997; Beghelli & Stowell 1997) and Hayashishita's (1999, 2000a, 2000b) theory augmented with Tomioka's (2004) proposal that Japanese *wh*+QPt quantifiers such as *dono N-mo* 'every N' cannot take object-wide scope due to their anti-topic status), it was predicted that English-speaking learners of Japanese could overcome poverty of the stimulus and acquire Japanese scope rigidity in SOV QP-QP sentences. The L2 data were compatible with this prediction and therefore with the two theories. Under the theories of Aoun & Li (1993), Hoji (1985) and Hornstein (1995), it was predicted that acquisition of scope rigidity would not be possible for English-speaking learners. This prediction was not supported by the experimental data. Regarding *wh*-QP interpretation, only Aoun & Li's (1993) theory addressed the relevant cross-linguistic differences. The L2 data on *wh*-QP interpretation, which showed that the lack of a pair-list interpretation in scrambled Japanese [*wh*-object...QP-subject...] questions can be acquired under poverty of the stimulus, do not support or undermine Aoun & Li's (1993) theory. However, it was argued that evidence of the availability of pair-list interpretations in native Korean poses a problem for the association in Aoun & Li's theory of 'question particle languages' (such as Japanese and Korean) with a lack of pair-list interpretations.

Section 6.3 briefly reiterated the aims of the methodology adopted, namely Schwartz & Sprouse's (2000) framework of a comparative interlanguage study incorporating L2 poverty of the stimulus. It was shown that the present study successfully met these aims and that the framework facilitates the formulation of useful conclusions about the roles of L1 transfer and UG for L2 acquisition theory. Specifically, comparison of learners whose L1s were English, Chinese or Korean showed that the divergence between the learners' developmental paths with respect to quantifier scope interpretation in Japanese reflected precisely the divergence between the three L1s with respect to the equivalent scope phenomena. This provides strong evidence of L1 transfer in L2 acquisition. In addition, evidence that two L2 poverty-of-the-stimulus problems were overcome by advanced learners of Japanese—the acquisition of scope rigidity in Japanese SOV QP-QP sentences by advanced English-speaking learners; and the acquisition of the absence of pair-list interpretations in Japanese scrambled [*wh*-object...QP-subject...] questions by advanced English-speaking, Chinese-speaking, and Korean-speaking learners—provides strong evidence that L2 acquisition is guided by UG.

7. Summary: main findings and conclusions

The main goal of this thesis was to investigate the L2 acquisition research questions given in Chapter 1, (4) and (5), repeated below in (248) and (249).

248. When a target language phenomenon *P* represents an L2 poverty-of-the-stimulus problem, are L2ers able to overcome the problem and acquire *P*?

249. Do L2ers show divergence with respect to *P* when their L1s are typologically distinct with respect to *P*?

An affirmative answer to (248) indicates that L2 acquisition is guided by UG, since, logically, neither the input nor the L1 can provide direct evidence about the facts about *P* under poverty of the stimulus. An affirmative answer to (249) indicates that L1 knowledge influences the L2 acquisition process. Investigation of (248) and (249) was conducted by means of a quantitative experimental study of two quantifier scope interpretation phenomena in L2 Japanese. The hypotheses are summarised in (250) and (251).[211]

250. *L2 Poverty-of-the-stimulus*
   a. Advanced learners of Japanese whose L1 is English will (correctly) reject non-native-like object-wide scope in Japanese SOV QP-QP sentences.
   b. Advanced learners of Japanese whose L1 is English, Chinese, or Korean will (correctly) reject non-native-like pair-list interpretations of scrambled Japanese [*wh*-object...QP-subject...] questions.

251. *L1 transfer*
   a. Lower-proficiency learners of Japanese whose L1 is English will (incorrectly) accept non-native-like object-wide scope in Japanese SOV QP-QP sentences. By contrast, lower-proficiency L2 learners of Japanese whose L1 is Chinese or Korean will (correctly) reject non-native-like object-wide scope in Japanese SOV QP-QP sentences.
   b. i. Lower-proficiency learners of Japanese whose L1 is Chinese, English or Korean will (incorrectly) accept non-native-like pair-list answers to scrambled Japanese [*wh*-object...QP-subject...] questions.
      ii. In addition, the lower proficiency L1-Chinese-speaking learners will (incorrectly) reject individual answers to scrambled Japanese [*wh*-object...QP-subject...] questions.

---

[211] The hypotheses about *wh*-QP interpretation (250b & 251b) reflect revisions made in Chapter 5 on the basis of the findings from the native control data that (contra the claims of the theoretical literature) (i) pair-list answers to scrambled [*wh*-object...QP-subject...] questions are accepted in native Korean, and (ii) individual answers to questions with a QP-subject and *wh*-object are rejected in native Chinese.

An acceptability judgement task was developed to collect the relevant data. Six groups of L2 learners of Japanese took part in the study: intermediate and advanced groups with English, Chinese or Korean as their L1. In addition, there were four native control groups: Japanese, English, Chinese, and Korean.

The hypotheses in (250) were confirmed by the judgement data: some of the advanced English-speaking learners of Japanese consistently rejected non-native-like object-wide scope in Japanese SOV QP-QP sentences; and some of the advanced L1 English-, L1 Chinese- and L1 Korean-speaking learners of Japanese consistently rejected non-native-like pair-list interpretations of scrambled Japanese [*wh*-object...QP-subject...] questions. In both tasks, the group rates of native-like judgements were higher among the advanced groups than the rates of native-like judgements among the corresponding intermediate learner groups. This result was argued to provide evidence that L2 acquisition can take place under poverty of the stimulus. This, in turn, supports the theory that L2 acquisition is guided by UG.

Hypothesis (251a) was also confirmed: the intermediate-level English-speaking learners of Japanese accepted object-wide scope in Japanese SOV QP-QP sentences, while the intermediate Chinese-speaking and Korean-speaking learners rejected it. This was argued to provide evidence of L1 transfer in L2 acquisition, since object-wide scope is acceptable in native English, but not in native Korean or Chinese (in [S…O…] QP-QP sentences). Hypothesis (251b.i) was also confirmed: the intermediate-level learners of all three L1 backgrounds accepted non-native-like pair-list answers to scrambled Japanese [*wh*-object...QP-subject...] questions, as predicted by L1 transfer. Hypothesis (251b.ii) was not confirmed: the intermediate L1 Chinese-speaking learners did not reject individual answers; contra the prediction, they accepted them. However, this group's rate of acceptance of individual answers was depressed compared with the other two intermediate groups'. This divergence between the intermediate groups was argued to indicate L1 influence, even though the hypothesis was not confirmed: the depressed rate of acceptance of individual answers by the intermediate Chinese-speaking learner group reflects the rejection of individual answers in native Chinese, while the high rates of acceptance of individual answers by the intermediate English-speaking and Korean-speaking groups reflects the high acceptability of individual answers in native English and native Korean.

Appendix 1: Phase 1 test material

About this appendix

This appendix relates to Phase 1 of the experimental work, and should be used in conjunction with Chapter 4. The appendix comprises two sections: Appendix 1A lists the test sentences used in Phase 1; Appendix 1B shows the corresponding pictures.

Note that, in Appendix 1A, each test item is identified by an index comprising two numbers and a letter. The first number, 1–10, indicates the sentence type (from Types 1–10; see Chapter 4). Each type has up to four exemplar sentences, and these are indicated by the second number, 1–4. The letter, a–d, indicates the word order and scope condition of that item ('a' = SOV order, subject-wide scope; 'b' = SOV order, object-wide scope; 'c' = OSV order, subject-wide scope; 'd' = OSV order, object-wide scope). These indices are used in Appendix 1B to identify which picture corresponds to which test item.

Appendix 1A: Test sentences

Notes:

1. As described in Chapter 4, the test battery was divided into two sets. The item number within each set is indicated for each test item in the 'Set 1' and 'Set 2' columns, below.
2. For translations of the scrambled items (those with 'c' and 'd' indices), see the preceding non-scrambled ('a' and 'b') items.

index	sentence	item number set 1	item number set 2
1.1a	Dareka-ga dono hon-mo yonde-iru.		50
1.1b	someone-NOM every book-QPt is reading 'Someone is reading every book.'	40	
1.1c		20	
1.1d	Dono hon-mo dareka-ga yonde-iru.		73
1.2a	Dareka-ga kono resutoran-no dono ryouri-mo tabete-mita.	31	
1.2b	someone-NOM this restaurant-GEN every dish-QPt eat-tried 'Someone tried every dish of this restaurant.'		69
1.2c			16
1.2d	Kono resutoran-no dono ryouri-mo dareka-ga tabete-mita.	28	
1.3a	Dareka-ga dono neko-mo nadete-iru.	64	
1.3b	someone-NOM every cat-QPt is stroking 'Someone is stroking every cat.'	14	
1.4a	Dareka-ga dono bousi-mo kabutte-mita.		26
1.4b	someone-NOM every hat-QPt wear-tried 'Someone tried on every hat.'		49
1.4c		25	
1.4d	Dono bousi-mo dareka-ga kabutte-mita.	57	
2.1a	Dareka-ga subete-no hon-o yonde-iru.	52	

index	sentence	item number set 1	item number set 2
2.1b	someone-NOM all-GEN book-ACC is reading 'Someone is reading all the books.'		22
2.1c	Subete-no hon-o dareka-ga yonde-iru.		39
2.1d		47	
2.2a	Dareka-ga kono resutoran-no subete-no ryouri-o tabete-mita.		59
2.2b	someone-NOM this restaurant-GEN all-GEN dish-ACC eat-tried 'Someone tried all the dishes of this restaurant.'	63	
2.2c	Kono resutoran-no subete-no ryouri-o dareka-ga tabete-mita.	4	
2.2d			28
2.3a	Dareka-ga subete-no neko-o nadete-iru.		51
2.3b	someone-NOM all-GEN cat-ACC is stroking 'Someone is stroking all the cats.'		54
2.3c	Subete-no neko-o dareka-ga nadete-iru.	18	
2.3d		53	
2.4a	Dareka-ga subete-no bousi-o kabutte-mita.	66	
2.4b	someone-NOM all-GEN hat-ACC wear-tried 'Someone tried on all the hats.'	22	
3.1a	Hutari-no otokonoko-ga dono teeburu-mo huita.	23	
3.1b	two.HUMAN-GEN boys-NOM every table-QPt wiped 'Two boys wiped every table.'		21
3.1c	Dono teeburu-mo hutari-no otokonoko-ga huita.		14
3.1d		13	
3.2a	Hutari-no otoko-ga dono doa-mo nutta.		2
3.2b	two.HUMAN-GEN men-NOM every door-QPt painted. 'Two men painted every door.'	35	
3.3a	Hutari-no onnanoko-ga dono mado-mo aratta.	27	
3.3b	two.HUMAN-GEN girls-NOM every window-QPt washed 'Two girls washed every window.'	30	
3.3c	Dono mado-mo hutari-no onnanoko-ga aratta.		46
3.3d			60
3.4a	Hutari-no kangohu-ga dono kanzya-mo kanbyou-sita.		13
3.4b	two.HUMAN-GEN nurse-NOM every patient-QPt looked-after. 'Two nurses looked after every patient.'		24
3.4c	Dono kanzya-mo hutari-no kangohu-ga kanbyou-sita.	16	
3.4d		5	
4.1a	Hutari-no otokonoko-ga subete-no teeburu-o huita.	42	
4.1b	two.HUMAN-GEN boys-NOM all-GEN table-acc wiped 'Two boys wiped all the tables.'	3	
4.1c	Subete-no teeburu-o hutari-no otokonoko-ga huita.		25
4.1d			53
4.2a	Hutari-no otoko-ga subete-no doa-o nutta.		12
4.2b	two.HUMAN-GEN men-NOM all-GEN door-acc painted. 'Two men painted all the doors.'		5
4.2c	Subete-no doa-o hutari-no otoko-ga nutta.	61	
4.2d		73	
4.3a	Hutari-no onnanoko-ga subete-no mado-o aratta.	8	
4.3b	two.HUMAN-GEN girls-NOM all-GEN window-acc washed 'Two girls washed all the windows.'		40
4.4a	Hutari-no kangohu-ga subete-no kanzya-o kanbyou-sita.		47
4.4b	two.HUMAN-GEN nurse-NOM all-GEN patient-acc looked-after. 'Two nurses looked after all the patients.'	58	
4.4c	Subete-no kanzya-o hutari-no kangohu-ga kanbyou-sita.	55	

index	sentence	item number set 1	item number set 2
4.4d			63
5.1a	Dareka-ga daremo-o aisite-iru. someone-nom everyone-acc loves 'Someone loves everyone.'		30
5.1b		56	
5.1c	Daremo-o dareka-ga aisite-iru.	9	
5.1d			35
5.2a	Dareka-ga daremo-o kiratte-iru. someone-nom everyone-acc hates 'Someone hates everyone.'	37	
5.2b			33
5.2c	Daremo-o dareka-ga kiratte-iru.		61
5.2d		51	
5.3a	Dareka-ga daremo-o mite-iru. someone-nom everyone-acc is-watching 'Someone is watching everyone.'	2	
5.3b		69	
5.3c	Daremo-o dareka-ga mite-iru.		38
5.3d			72
5.4a	Dareka-ga daremo-o sikatte-iru. someone-nom everyone-acc is-scolding 'Someone is scolding everyone.'		66
5.4b			56
6.1a	Dareka-ga minna-o aisite-iru. someone-nom everyone-acc loves 'Someone loves everyone.'	59	
6.1b			48
6.1c	Minna-o dareka-ga aisite-iru.		64
6.1d		72	
6.2a	Dareka-ga minna-o kiratte-iru. someone-nom everyone-acc hates 'Someone hates everyone.'		19
6.2b		45	
6.2c	Minna-o dareka-ga kiratte-iru.	1	
6.2d			42
6.3a	Dareka-ga minna-o mite-iru. someone-nom everyone-acc is-watching 'Someone is watching everyone.'		7
6.3b			43
6.4a	Dareka-ga minna-o sikatte-iru. someone-nom everyone-acc is-scolding 'Someone is scolding everyone.'	19	
6.4b		68	
6.4c	Minna-o dareka-ga sikatte-iru.		45
6.4d			6
7.1a	Dareka-ga dono isu-ni-mo suwatte-mita. someone-NOM every chair-LOC-QPt sit-tried 'Someone sat (*lit.* tried sitting) in every chair.'	6	
7.1b		10	
7.1c	Dono isu-ni-mo dareka-ga suwatte-mita.		18
7.1d			11
7.2a	Dareka-ga dono teeburu-no-ue-de-mo odotta. someone-NOM every table-GEN-top-LOC-QPt danced 'Someone danced on every table.'		10
7.2b			4
7.2c	Dono teeburu-no-ue-de-mo dareka-ga odotta.	24	
7.2d		60	
7.3a	Dareka-ga dono kuruma-ni-mo notte-mita. someone-NOM every car-LOC-QPt get in-tried 'Someone got in (*lit.* tried getting in) in every car.'	29	
7.3b			36
7.3c	Dono kuruma-ni-mo dareka-ga notte-mita.		15
7.3d		67	
7.4a	Dareka-ga dono mise-ni-mo haitta.		57

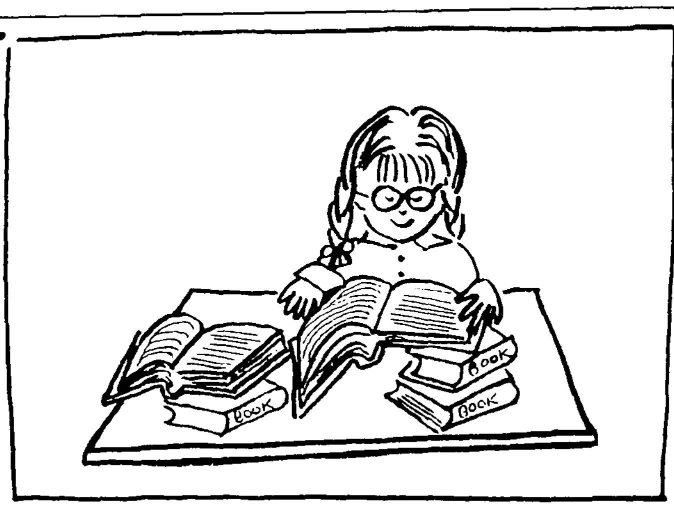
index	sentence	item number set 1	set 2
7.4b	someone-NOM every car-LOC-QPt entered 'Someone went into every shop.'	32	
8.1a	Ni-wa-no tori-ga dono ki-e-mo tonda.		44
8.1b	two-CL-GEN bird-NOM every tree-LOC-QPt flew 'Two birds flew to every tree.'		32
8.1c	Dono ki-e-mo ni-wa-no tori-ga tonda.	54	
8.1d		36	
8.2a	Hutari-no kankoukyaku-ga dono machi-ni-mo itta.	26	
8.2b	two.HUMAN-GEN tourist-NOM every city-LOC-QPt went 'Two tourists went to every city.'	38	
8.2c	Dono machi-ni-mo hutari-no kankoukyaku-ga itta.		29
8.2d			17
8.3a	San-biki-no neko-ga dono kussyon-ni-mo neta.	71	
8.3b	three-CL-GEN cat-NOM every cushion-LOC-QPt slept 'Three cats slept on every cushion.'	65	
8.4c	Dono kussyon-ni-mo san-biki-no neko-ga neta.		71
8.4d			34
9.1a	Dareka-ga dono heya-ni-mo ita.	15	
9.1b	someone-NOM every room-LOC-QPt was 'Someone was in every room.'	39	
9.1c	Dono heya-ni-mo dareka-ga ita.	44	
9.1d		33	
9.2a	Dareka-ga dono hoomu-ni-mo tuita.		37
9.2b	someone-NOM every platform-LOC-QPt arrived 'Someone arrived on every ticket platform.'		27
9.2c	Dono madoguti-ni-mo dareka-ga tuita.		52
9.2d			8
10.1a	Ni-mai-no hanabira-ga dono isi-ni-mo otita.	41	
10.1b	two-CL-GEN petal-NOM every rock-LOC-QPt fell 'Two petals fell onto every rock.'	7	
10.1c	Dono isi-ni-mo ni-mai-no hanabira-ga otita.		31
10.1d			23
10.2a	San-ko-no huusen-ga dono yane-e-mo agatta.		55
10.2b	two-CL-GEN balloon-NOM every roof-LOC-QPt floated 'Two balloons floated to every roof.'		65
10.2c	Dono yane-e-mo san-ko-no huusen-ga agatta.	21	
10.2d		12	

Distractor items ('D' in the index = 'Distractor'):

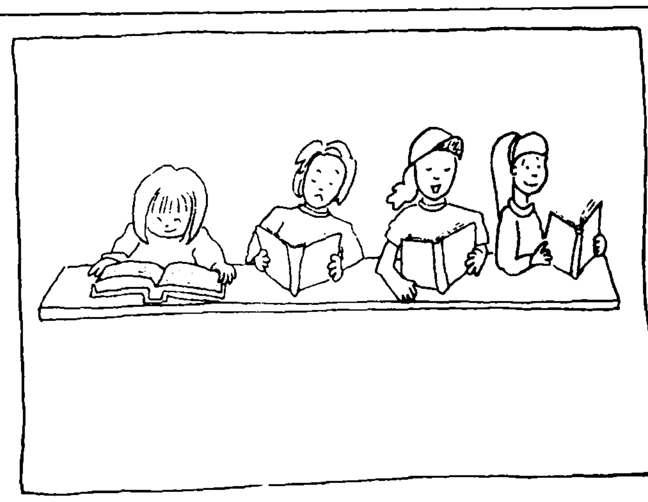
index	sentence	item number set 1	item number set 2
D1x	Hutari-no kodomo-ga Hanako-o ie-o deru toki ni mita. two.HUMAN-GEN child-NOM Hanako-ACC house-ACC leave time at saw 'Two children saw Hanako as they/she left the house.'	17	9
D1y	Hutari-no kodomo-ga Hanako-ga ie-o deru toki ni mita. two.HUMAN-GEN child-NOM Hanako-NOM house-ACC leave time at saw 'Two children saw Hanako as she left the house.'	46	70
D2x	Keikan-ga hutari-no dorobou-ga ginkou-o deru tokoro-o mita.	50	20
D2y	police officer-NOM two.HUMAN-GEN thief-NOM bank-ACC leave moment-ACC saw 'The police officer(s) saw two thieves$_i$ as they$_i$ left the bank.'	43	3
D3x	Dareka-ga Taroo-o hadaka-de mita.	48	62
D3y	Someone-NOM Taro-ACC naked saw 'Someone saw Taro naked.'	70	68
D4x	Zyiroo to Mariko-ga kinou deeto-o sita. Jiro and Mariko-NOM yesterday date-ACC did Jiro and Mariko went on a date yesterday.	49	41
D4y	Zyon to Merii-ga kinou deeto-o sita. John and Mary-NOM yesterday date-ACC did John and Mary went on a date yesterday.	62	58
D5x	Ni-hiki-no usagi-ga oziisan-o ninzin-o toru toki ni mita.	34	1
D5y	two.CL-GEN rabbit-NOM old man-ACC carrot-ACC dig-up time at saw 'Two rabbits saw the old man as they/he dug up carrots.'	11	67

Appendix 1B: Pictures

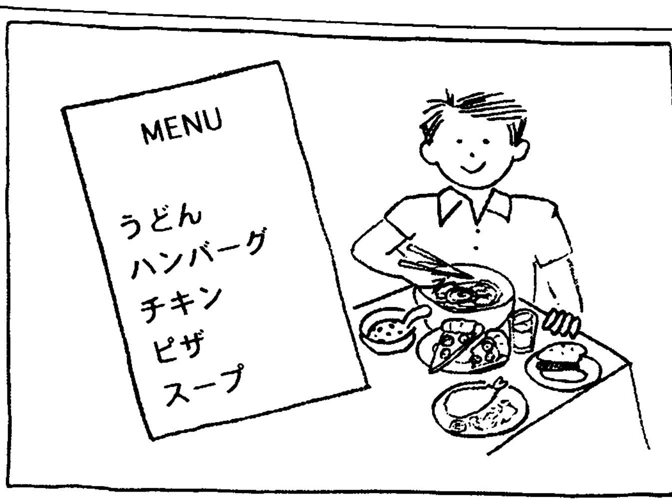
Note: A number of the pictures in the actual test were in colour. Please contact the author about viewing the pictures in colour, if necessary.



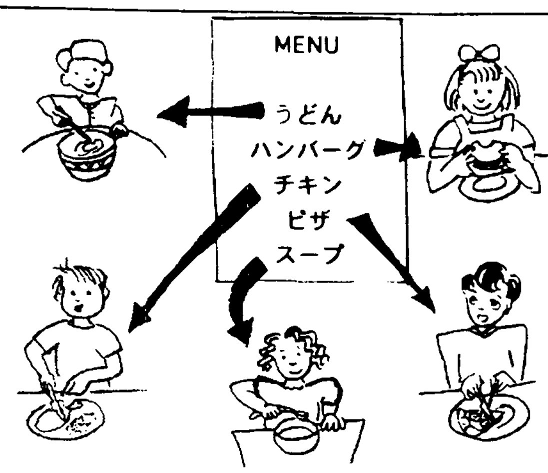
Picture for 1.1a, 1.1c, 2.1a, 2.1c.



Picture for 1.1b, 1.1d, 2.1b, 2.1d.

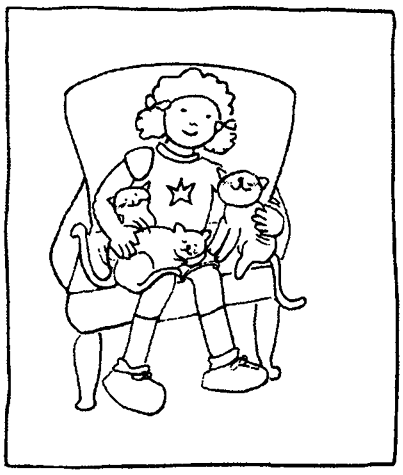


Picture for 1.2a, 1.2c, 2.2a, 2.2c.

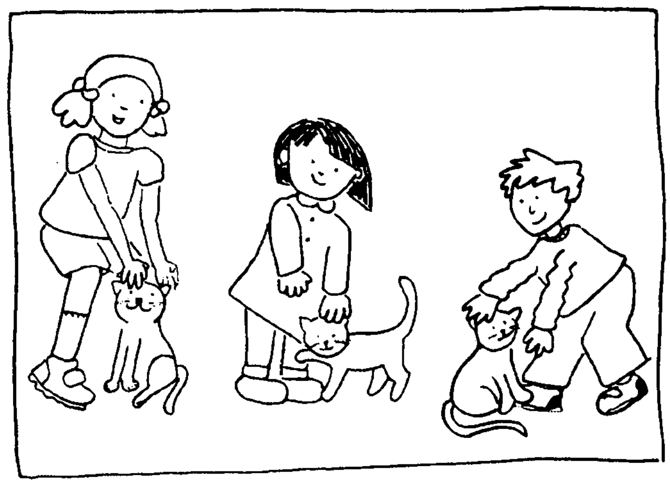


Picture for 1.1b, 1.1d, 2.2b, 2.2d.

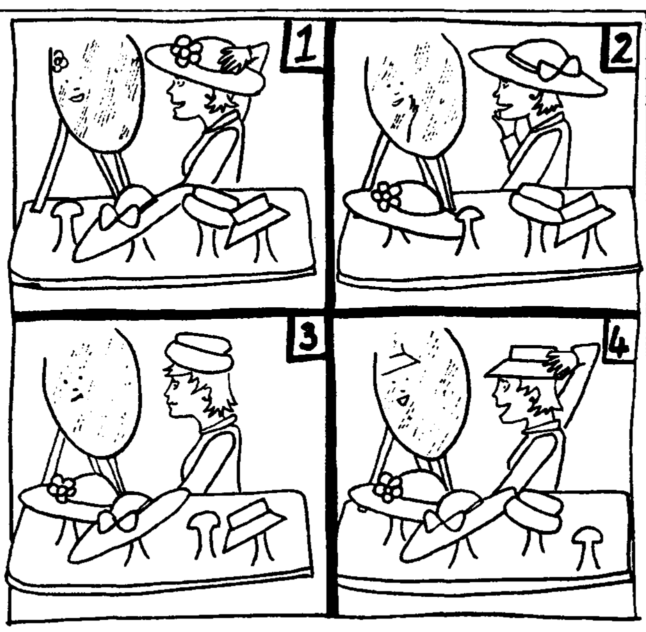
(**Translation of menu**: noodles, hamburger, chicken, pizza, soup)



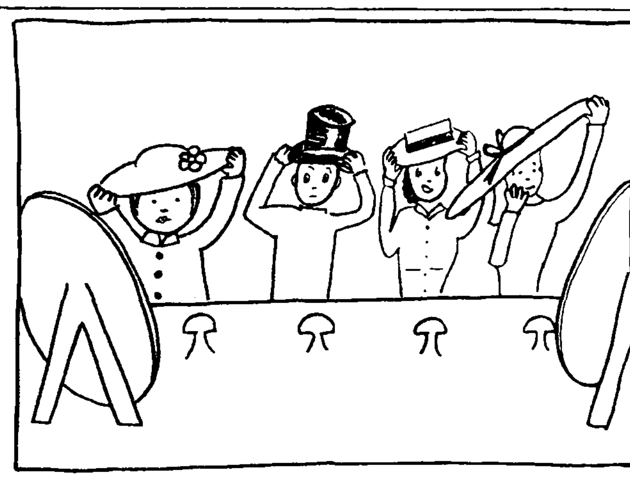
Picture for 1.3a, 2.3a, 2.3c.



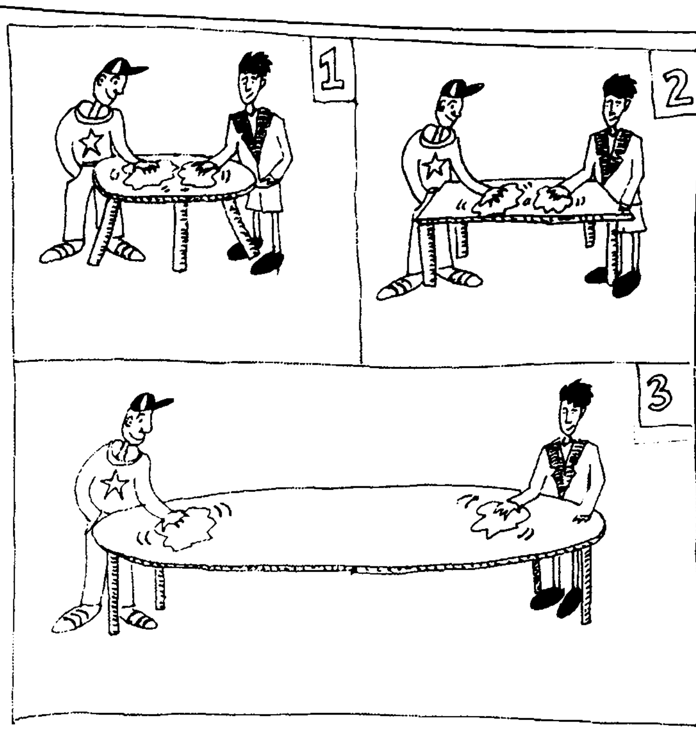
Picture for 1.3b, 2.3b, 2.3d.



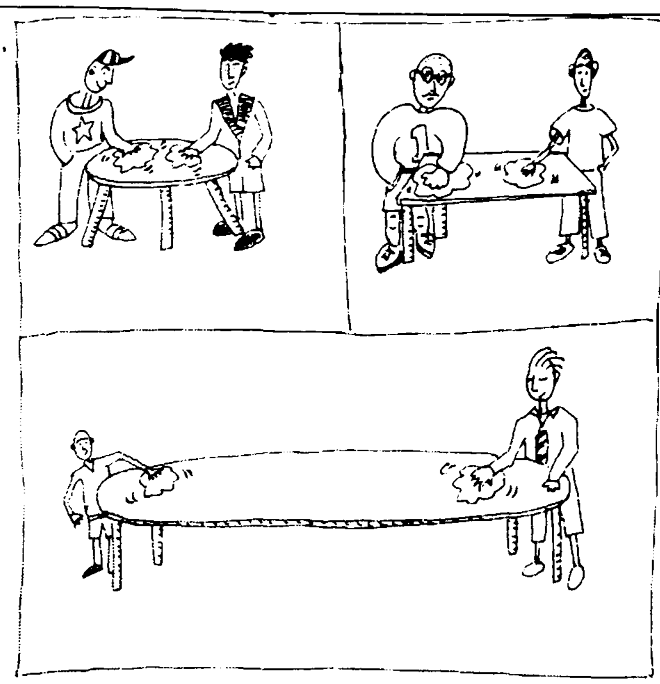
Picture for 1.4a, 1.4, 2.4a.



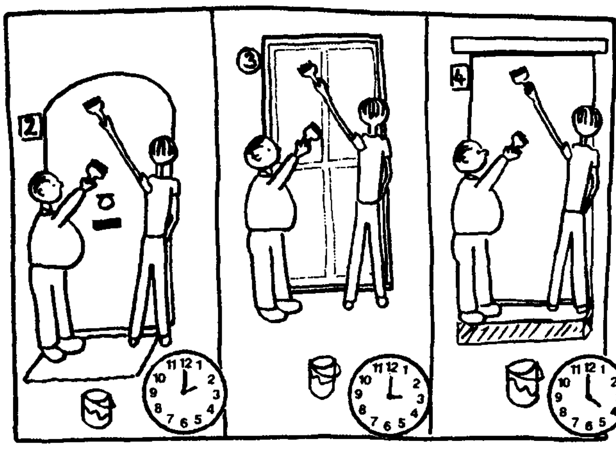
Picture for 1.1b, 1.1d, 2.4b.



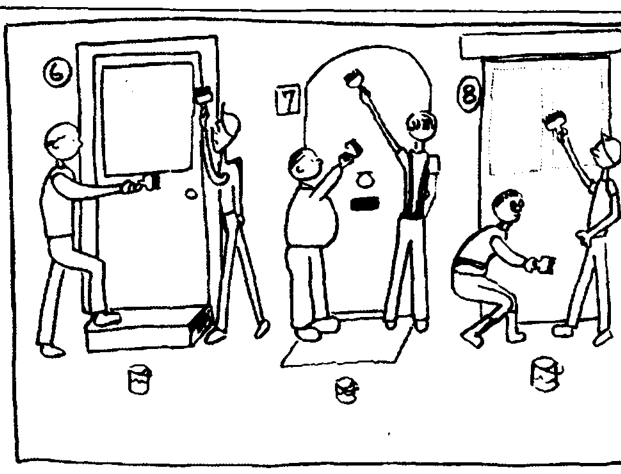
Picture for 3.1a, 3.1c, 4.1a, 4.1c.



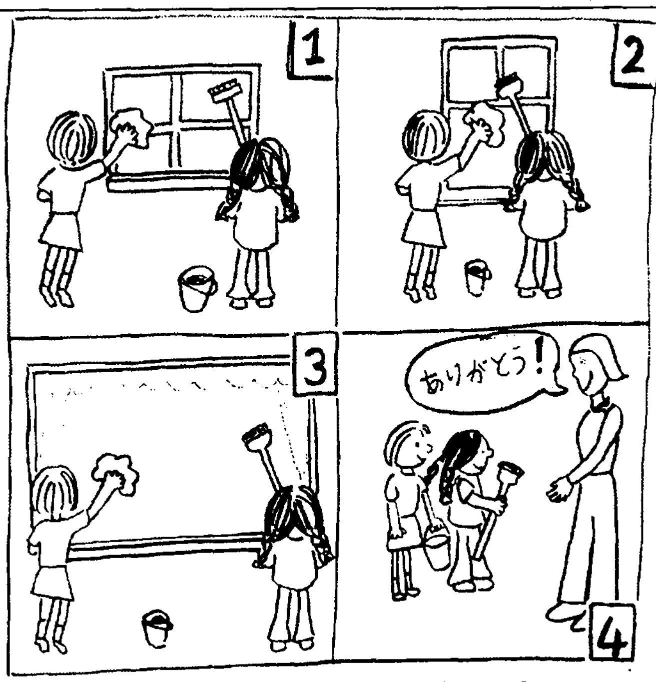
Picture for 3.1b, 3.1d, 4.1b, 4.1d.



Picture for 3.2a, 4.2a, 4.2c.

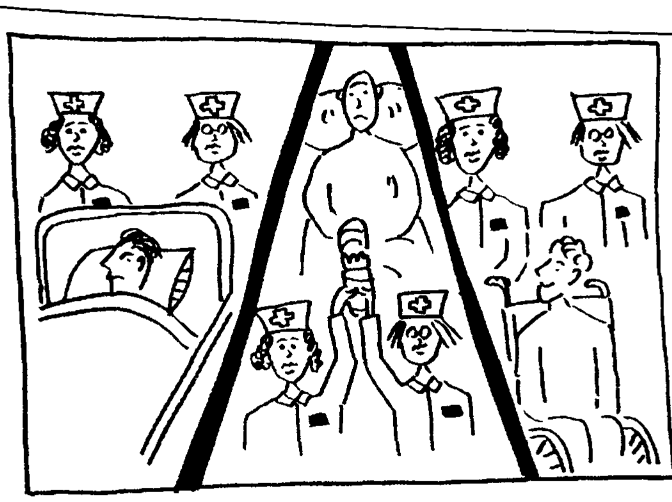


Picture for 3.2b, 4.2b, 4.2d.

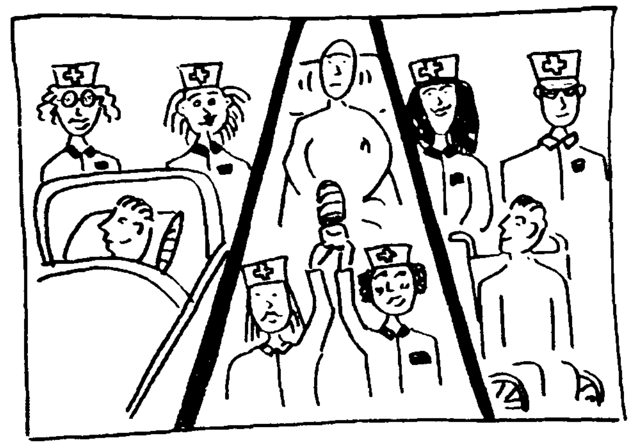


Picture for 3.3a, 3.3c, 4.3a.
(Translation of speech bubble: 'Thank you!')

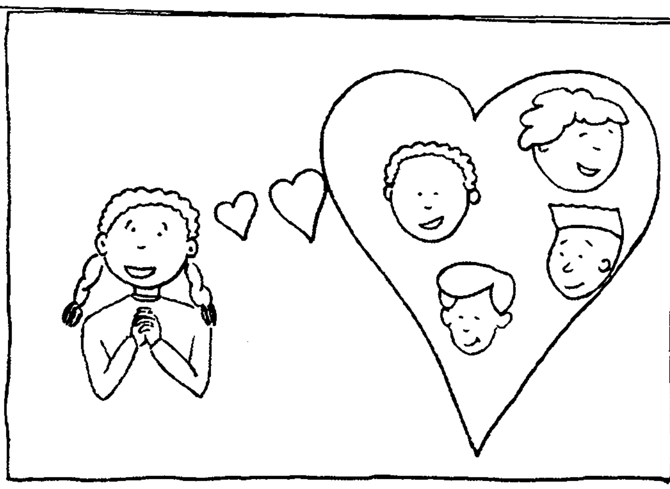
Picture for 3.3b, 3.3d, 4.3b.



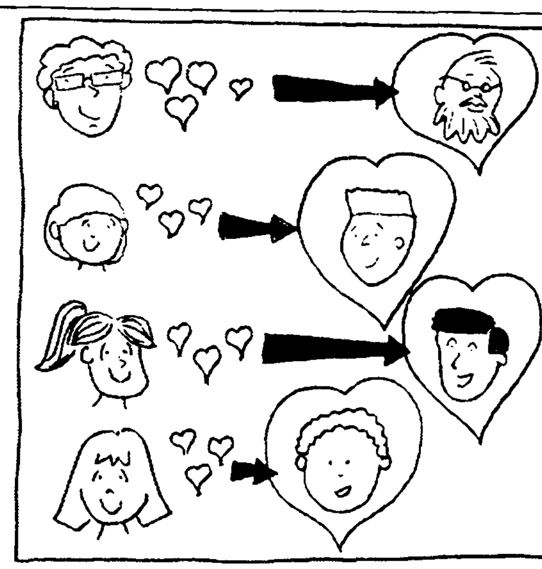
Picture for 3.4a, 3.4c, 4.4a, 4.4c.



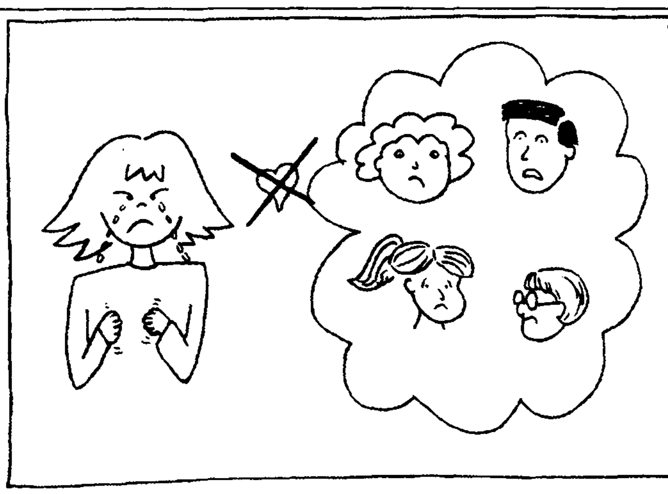
Picture for 3.4b, 3.4d, 4.4b, 4.4d.



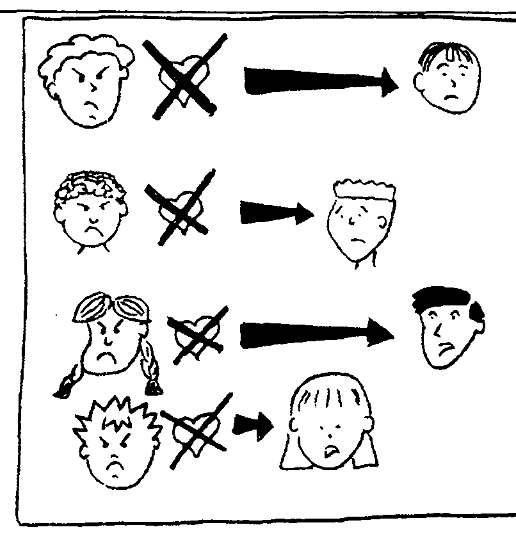
Picture for 5.1a, 5.1c, 6.1a, 6.1c.



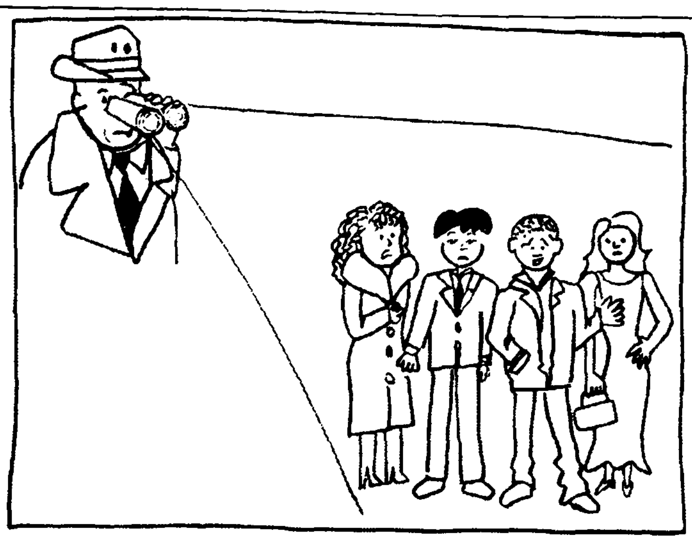
Picture for 5.1b, 5.1d, 6.1b, 6.1d.



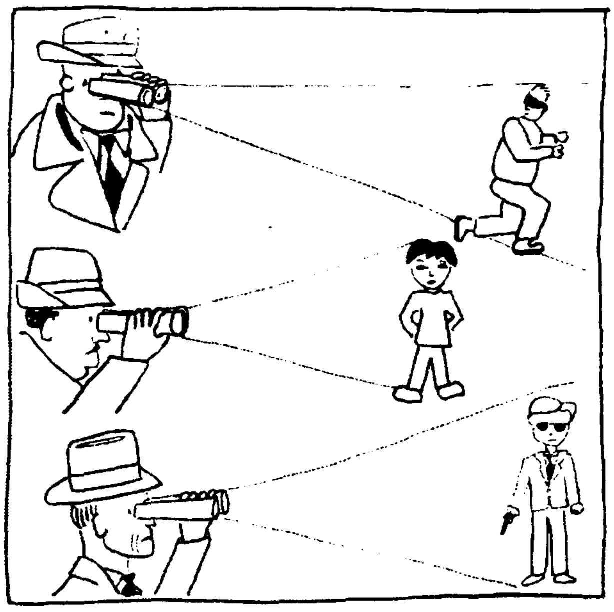
Picture for 5.2a, 5.2c, 6.2a, 6.2c.



Picture for 5.2b, 5.2d, 6.2b, 6.2d.



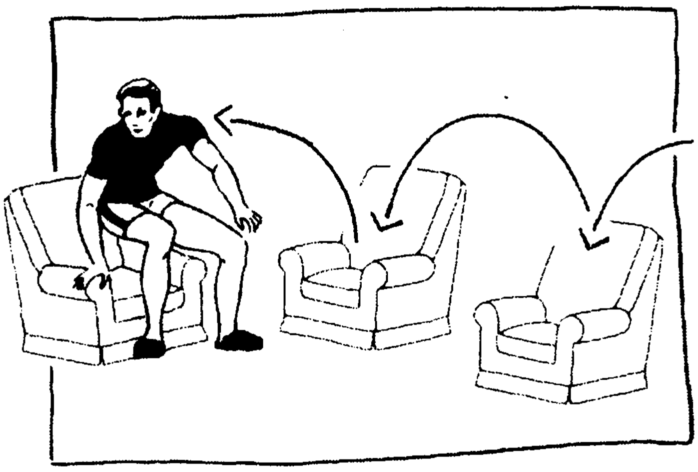
Picture for 5.3a, 5.3c, 6.3a.



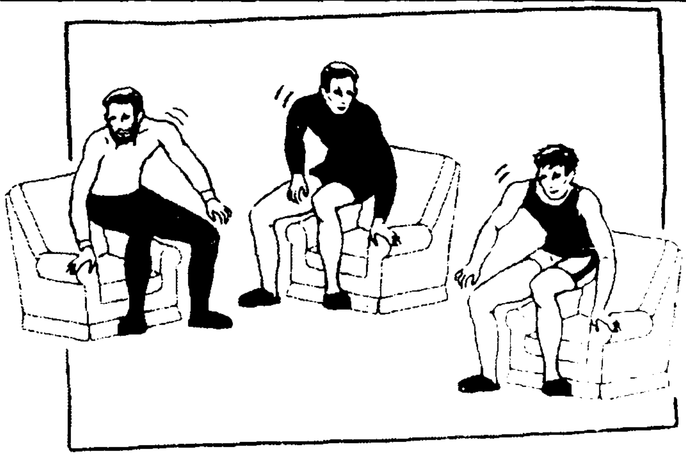
Picture for 5.3b, 5.3d, 6.3b.

Picture for 5.4a, 6.4a, 6.4c.

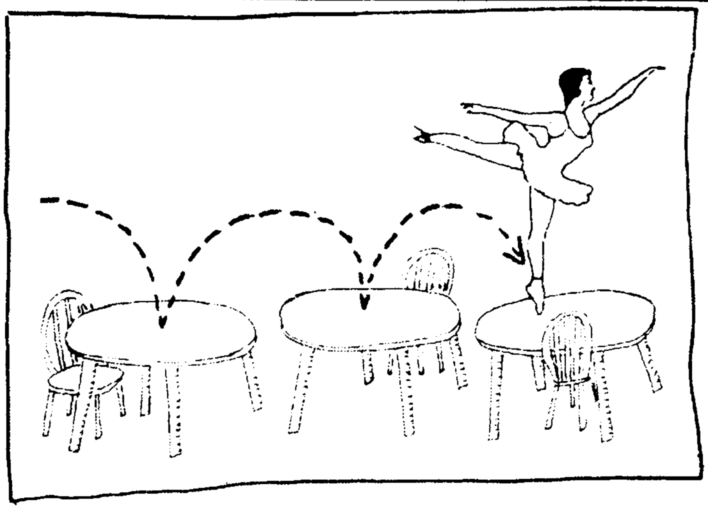
Picture for 5.4b, 6.4b, 6.4d.



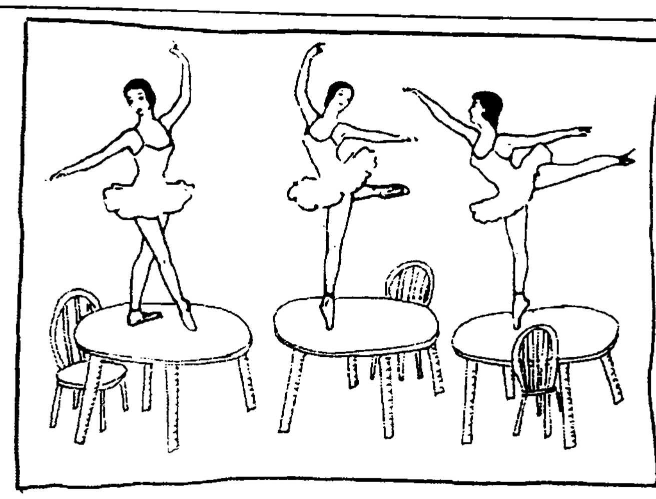
Picture for 7.1a, 7.1c.



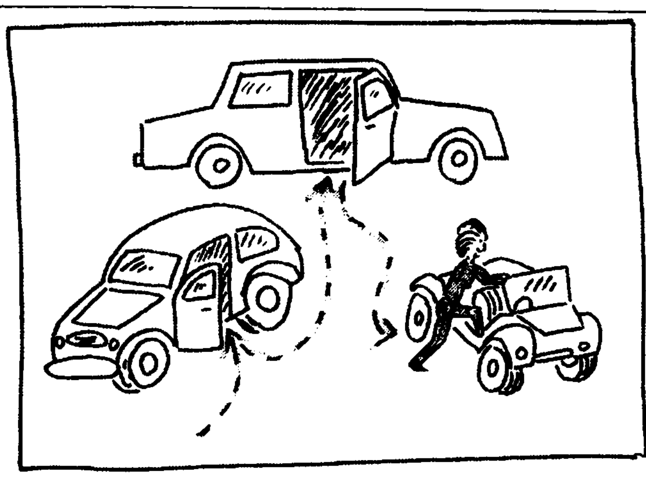
Picture for 7.1b, 7.1d.



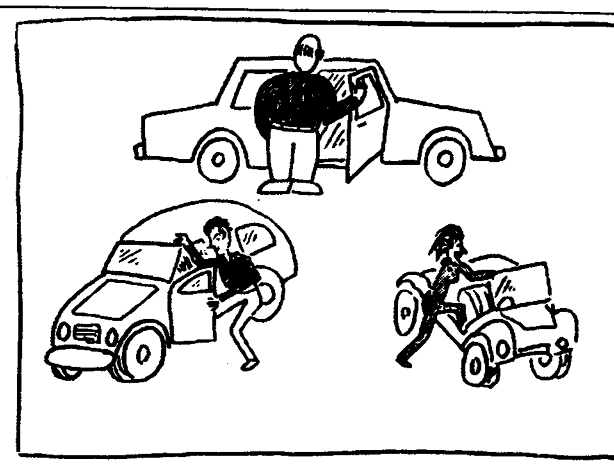
Picture for 7.2a, 7.2c.



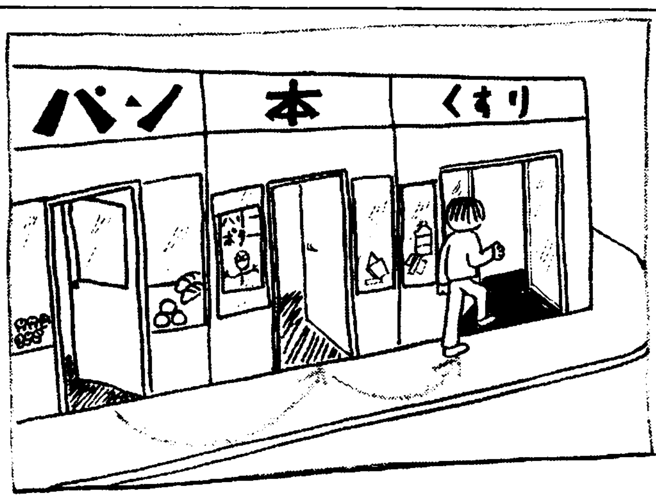
Picture for 7.2b, 7.2d.



Picture for 7.3a, 7.3c.

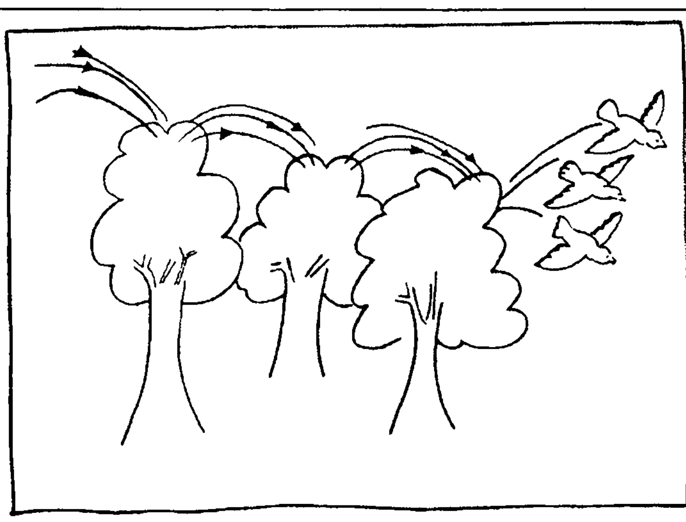


Picture for 7.3b, 7.3d.

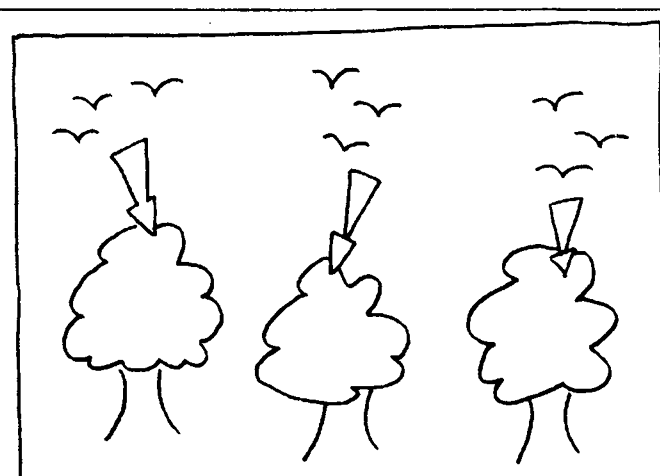


Picture for 7.4a.
(**Translation:** Bread, Books, Pharmacy)

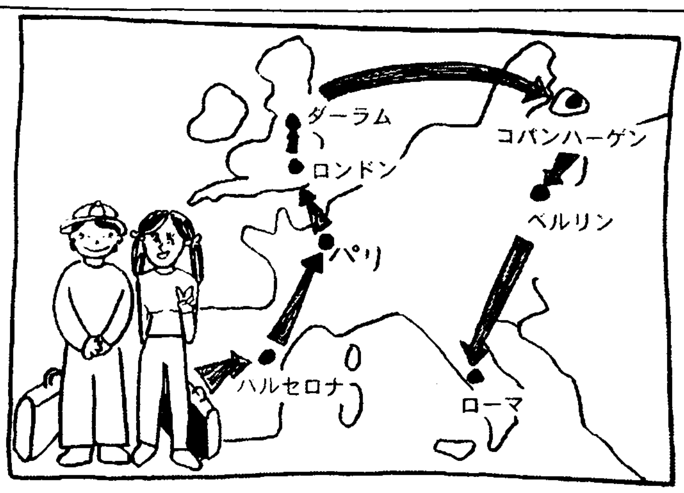
Picture for 7.4b.
(**Translation:** Supermarket, Liquor, Flowers)



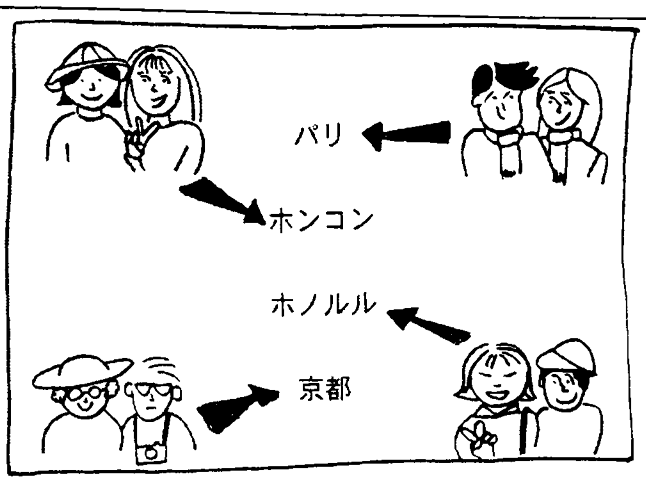
Picture for 8.1a, 8.1c.



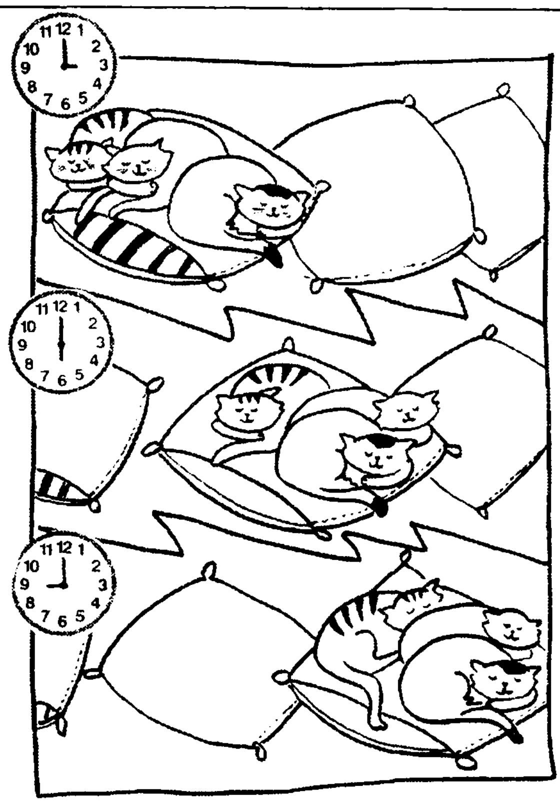
Picture for 8.1b, 8.1d.



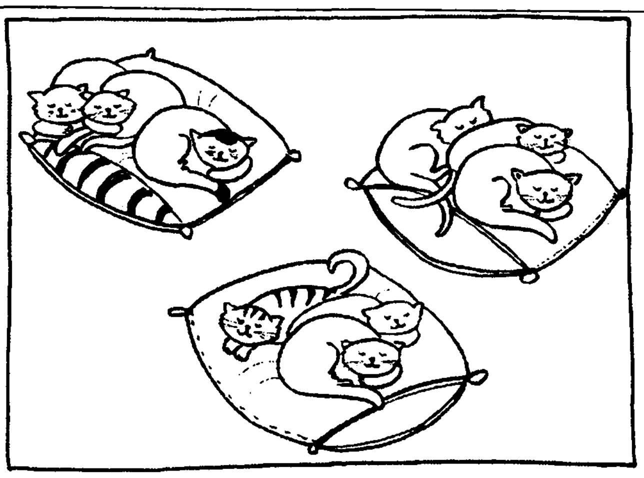
Picture for 8.2a, 8.2c.
(**Translation:** Barcelona, Paris. London, Durham, Copenhagen, Berlin, Rome)



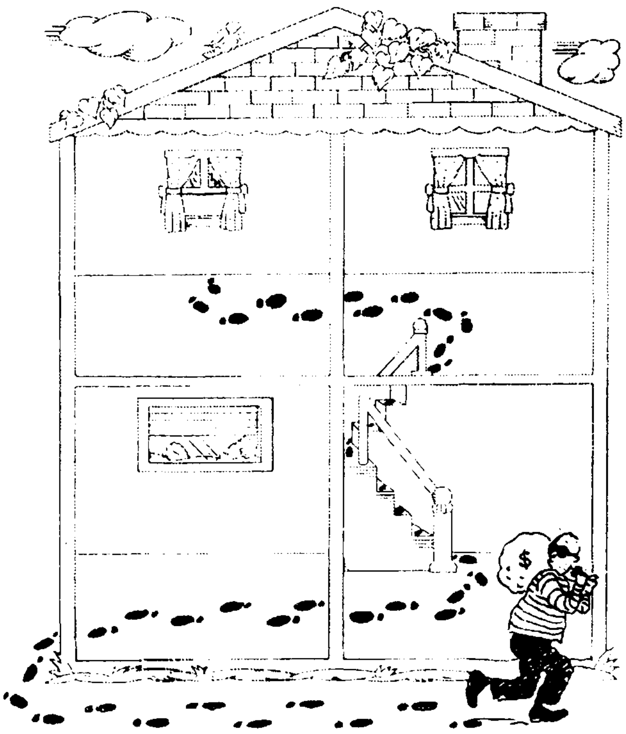
Picture for 8.2b, 8.2d.
(**Translation:** Paris, Hong Kong, Honolulu, Kyoto)



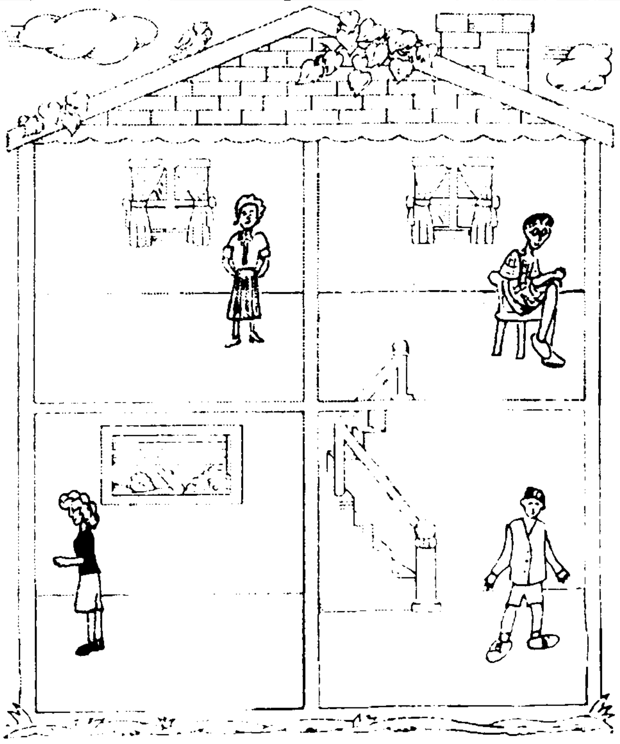
Picture for 8.3a, 8.3c.



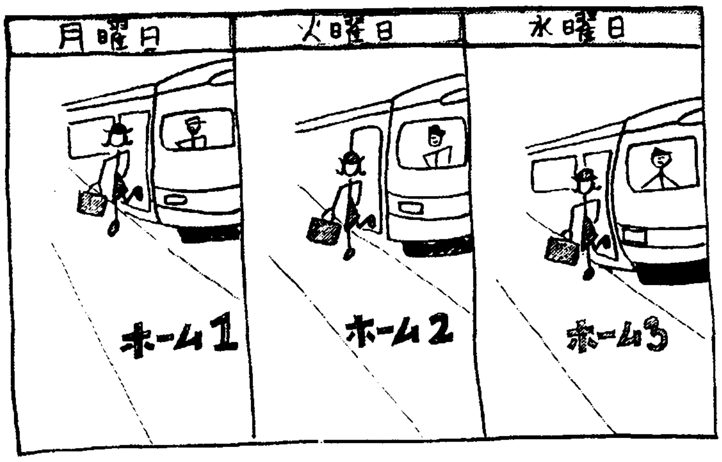
Picture for 8.3b, 8.3d.



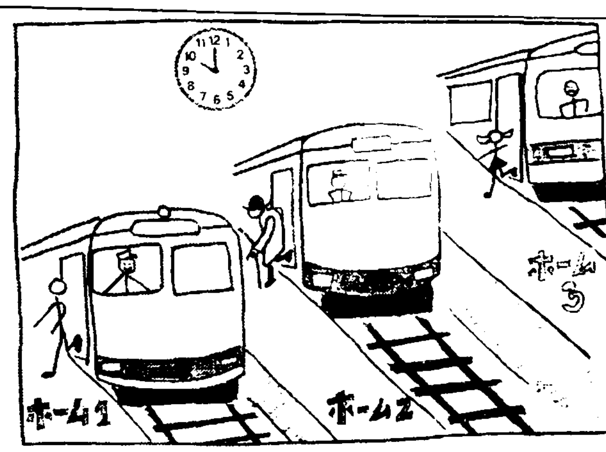
Picture for 9.1a, 9.1c.



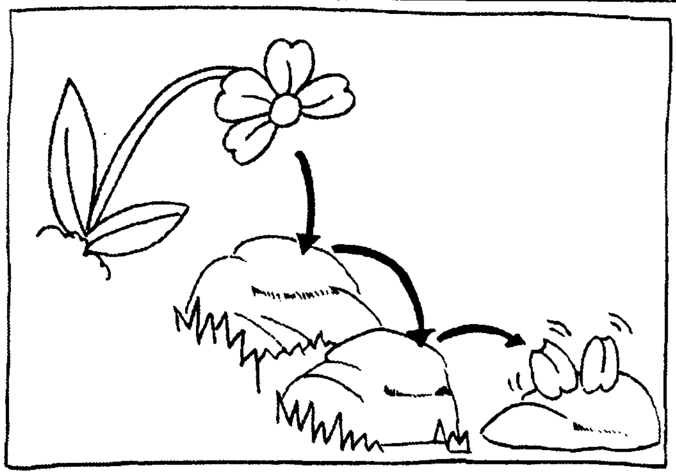
Picture for 9.1b, 9.1d.



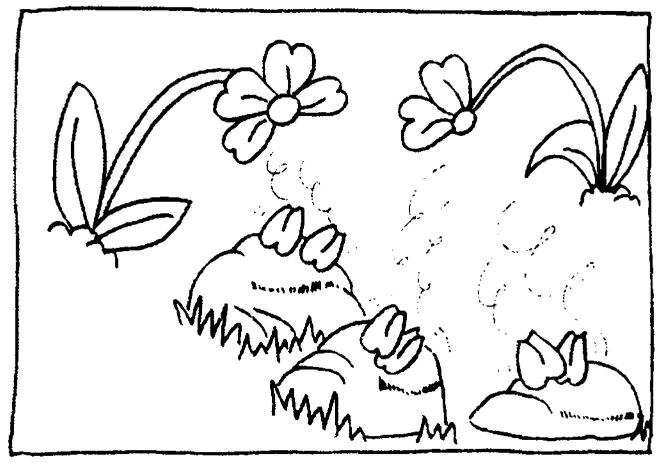
Picture for 9.2a, 9.2c.
(**Translation** (top): Monday, Tuesday, Wednesday; (bottom): Platform 1, Platform 2, Platform 3)



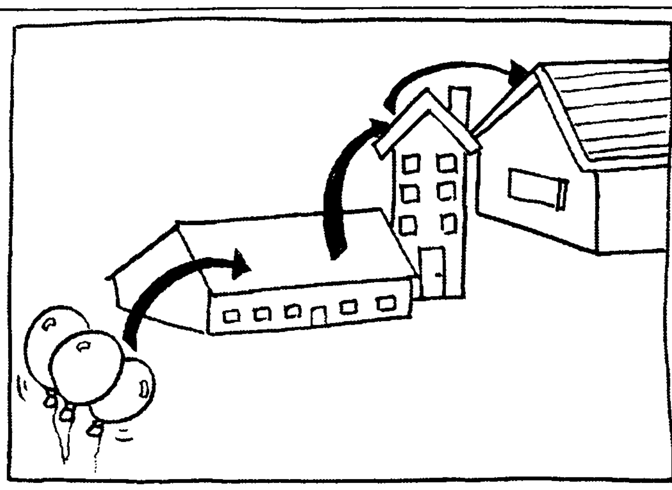
Picture for 9.2b, 9.2d.
(**Translation**: Platform 1, Platform 2, Platform 3)



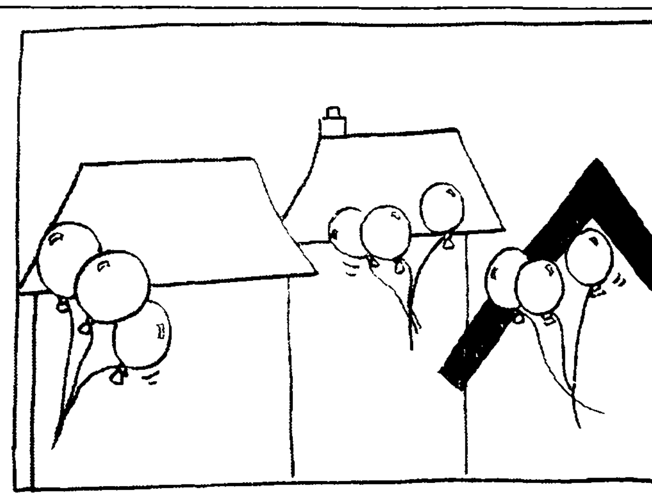
Picture for 10.1a, 10.1c.



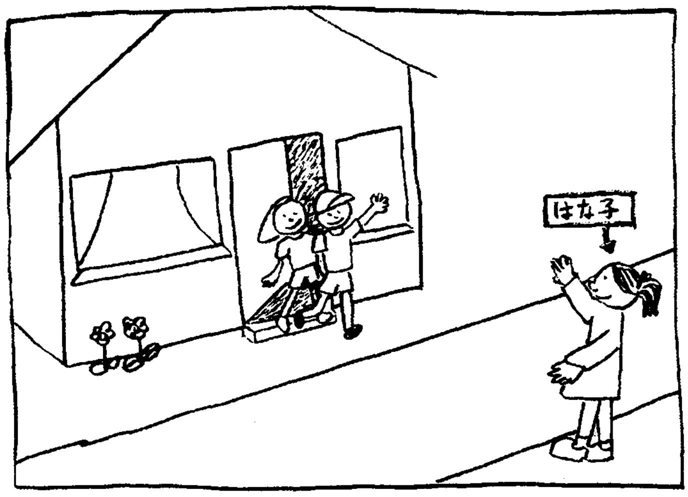
Picture for 10.1b, 10.1d.



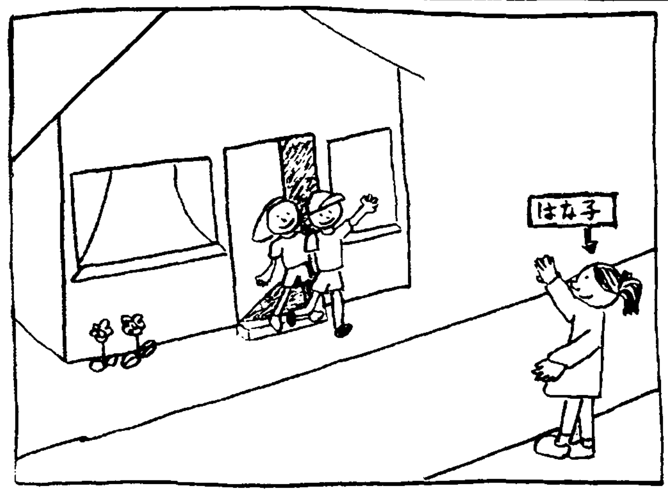
Picture for 10.2a, 10.2c.



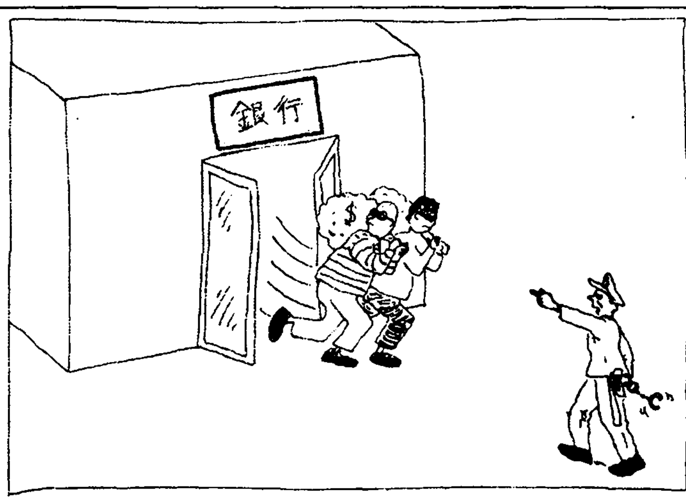
Picture for 10.2b, 10.2d.



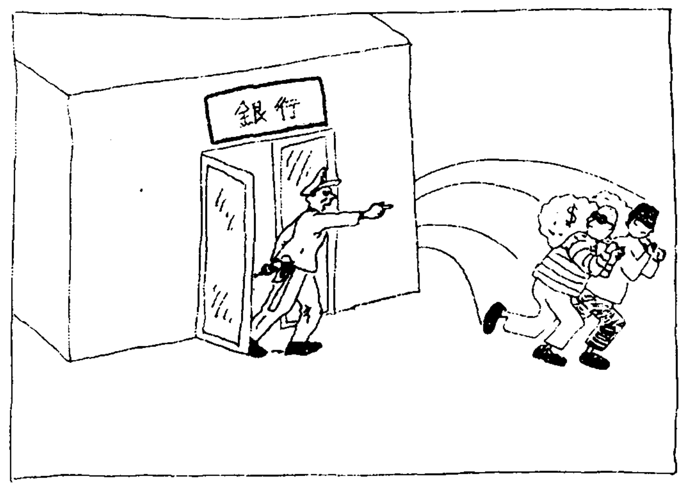
Picture for D1x.
(**Translation:** Hanako [girl's name])



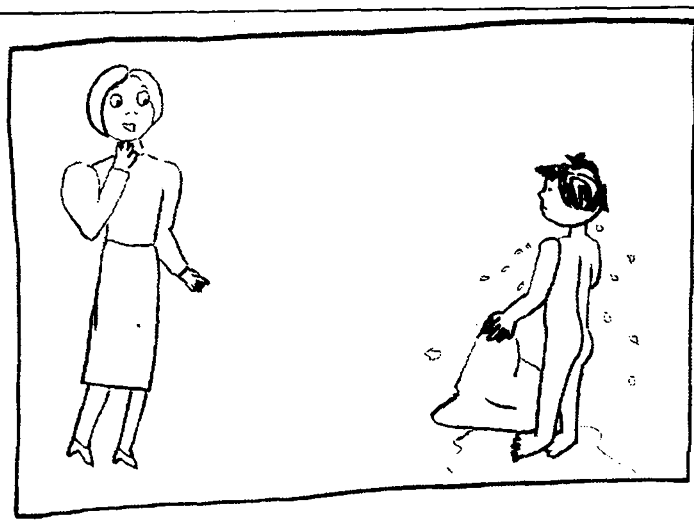
Picture for D1y.
(**Translation:** Hanako [girl's name])



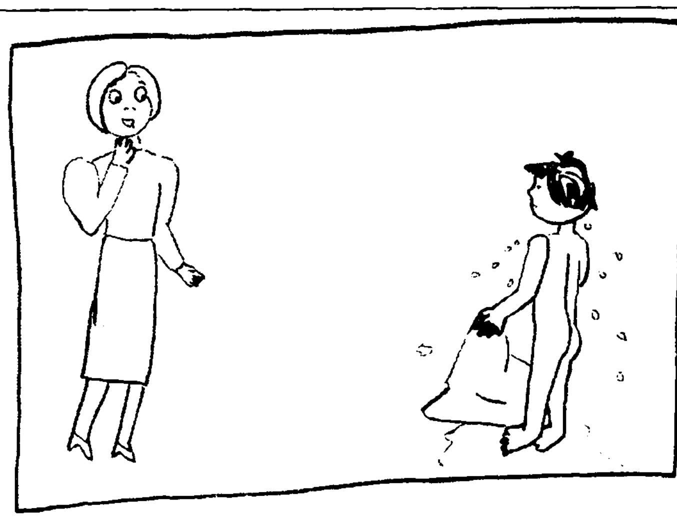
Picture for D2x.
(**Translation**: Bank)



Picture for D2y.
(**Translation**: Bank)

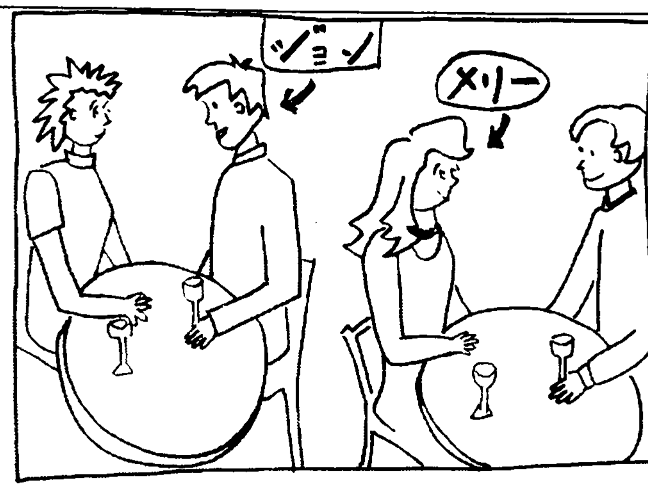


Picture for D3x.

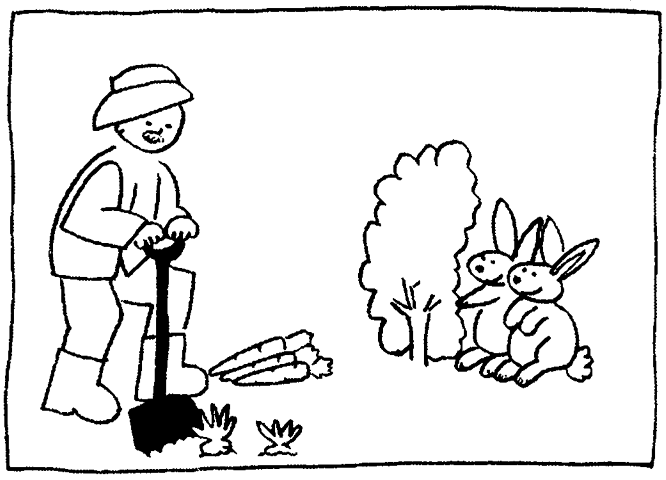


Picture for D3y.

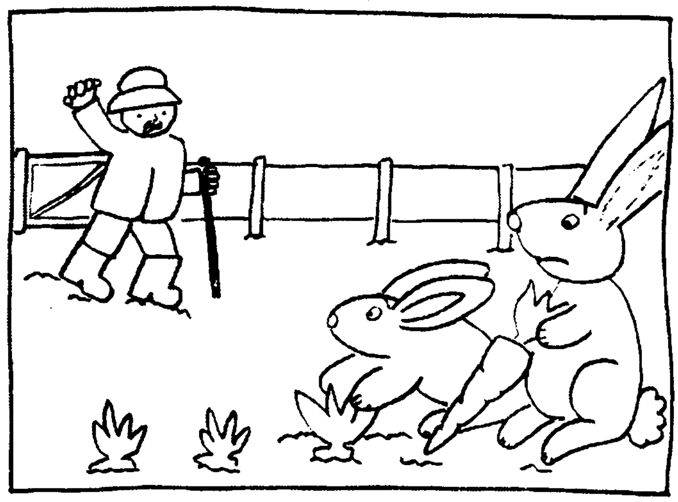
Picture for D4x.
(**Translation,** L–R: Mariko, Jiro)



Picture for D4y.
(**Translation,** L–R: John, Mary)



Picture for D5x.



Picture for D5y.

Appendix 2: Phase 2 test materials and results

About this appendix

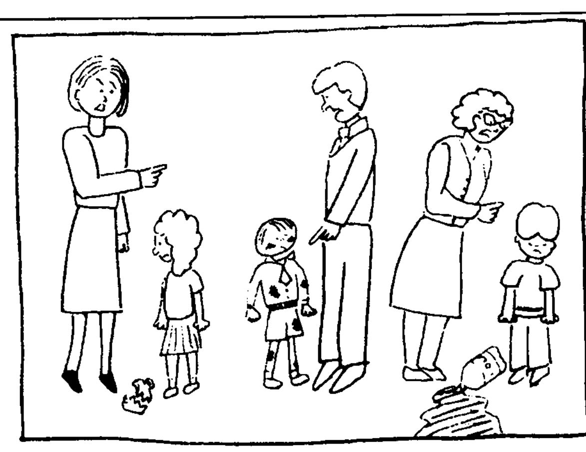
This appendix relates to Phase 2 of the experimental work, and should be used in conjunction with Chapter 4. The following appendices are provided:

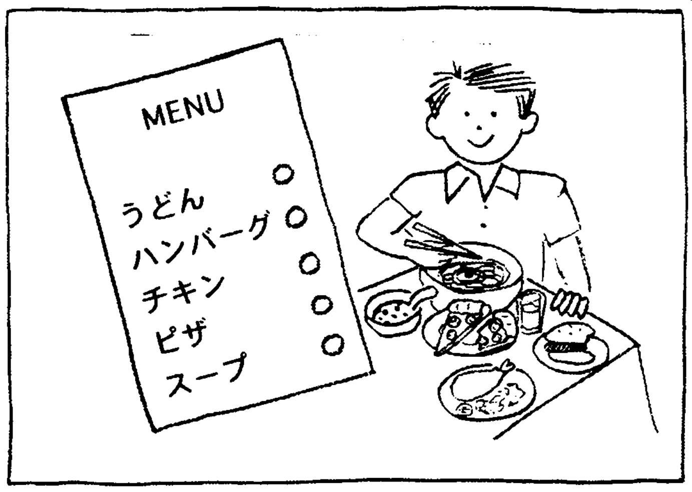
1. Appendix 2A: test sentences and pictures for Phase 2A
2. Appendix 2B: answer sheet used by learners of Japanese in Phase 2A
3. Appendix 2C: tables of results for Phase 2A
4. Appendix 2D: test sentences and written contexts for Phase 2B
5. Appendix 2E: tables of results for Phase 2B

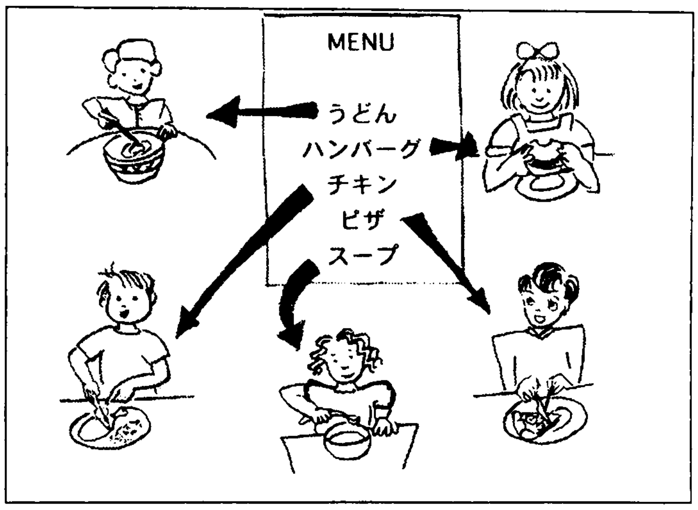
Appendix 2A: Test sentences and pictures

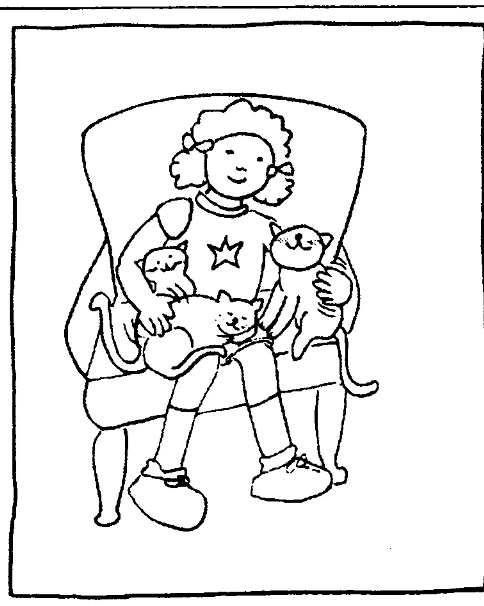
Notes:

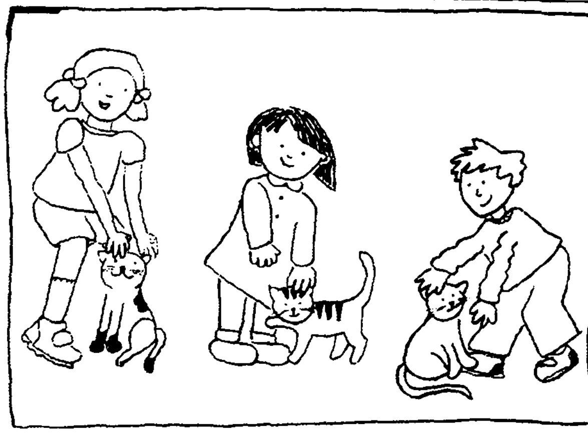
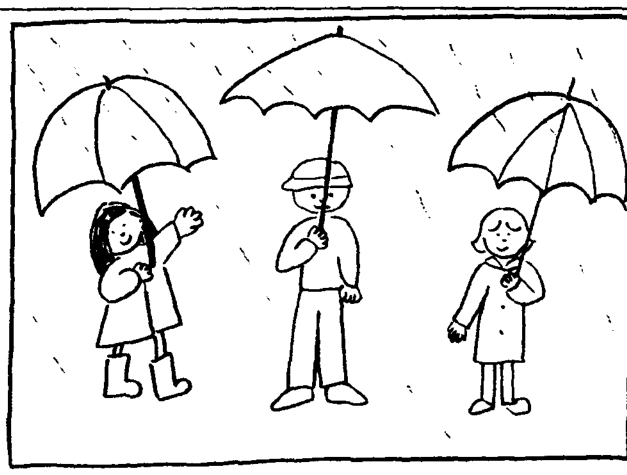
1. The English version of the test used the sentences given below as translations of the Japanese.
2. For convenience, English translations of the Japanese words appearing in the pictures are given below the pictures. These translations did not appear in the actual test.
3. In the English version of the test, all words in the pictures were in English.
4. Colour was used to facilitate comprehension of the pictures. Please contact the author about viewing the pictures in colour, if necessary.

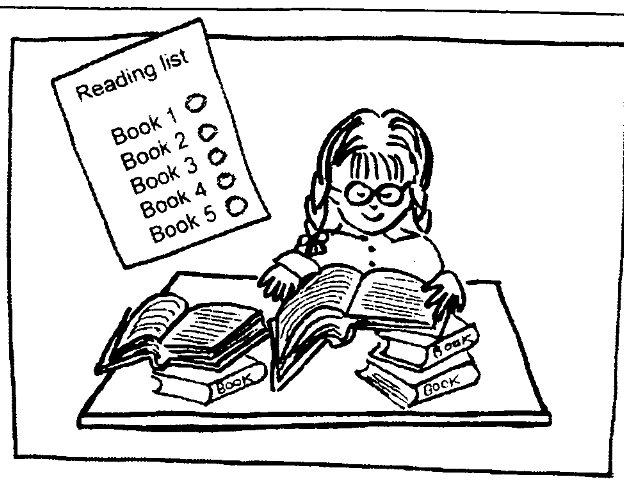
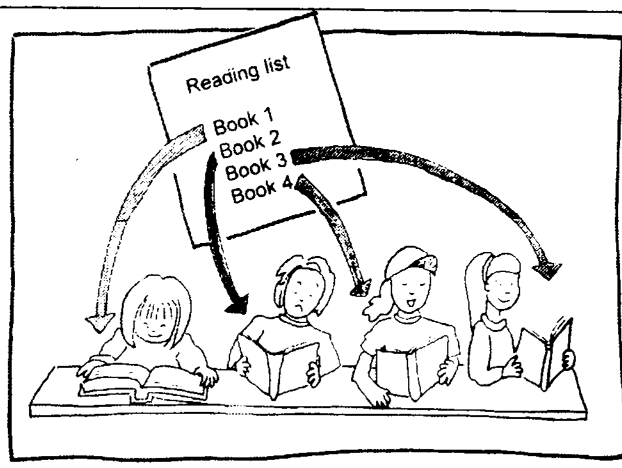
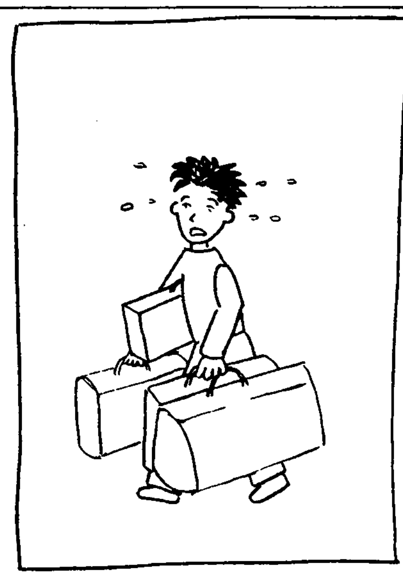
index	picture & sentence	item no.	
		order 1	order 2
1a	 だれかがどの子供も叱(しか)った。 Dareka-ga dono kodomo-mo sikatta. someone-NOM every child-QPt scolded 'Someone scolded every child.'	41	4
1b	 (*Sentence as 1a*)	29	18

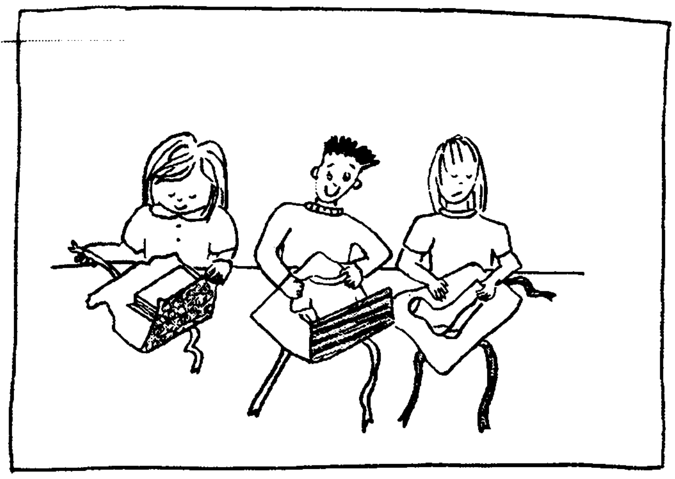
<table>
<tr><td>2a</td><td>

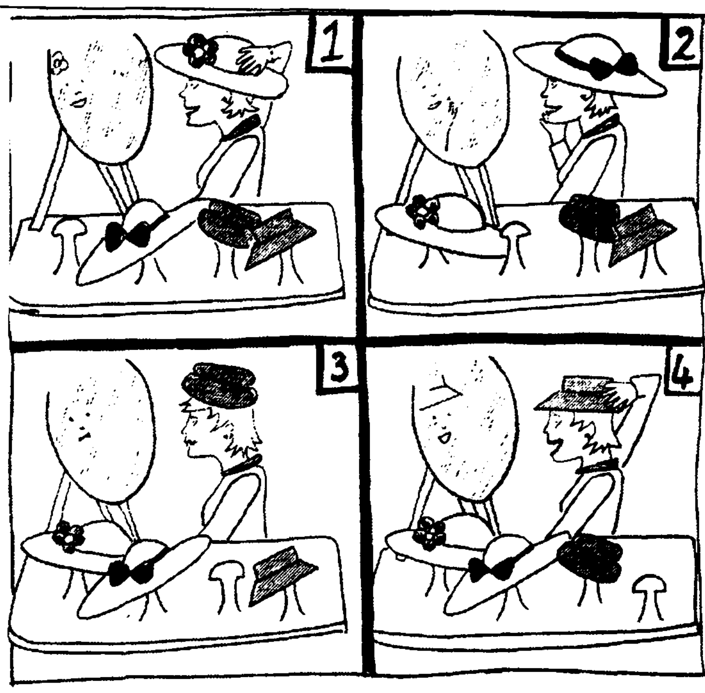
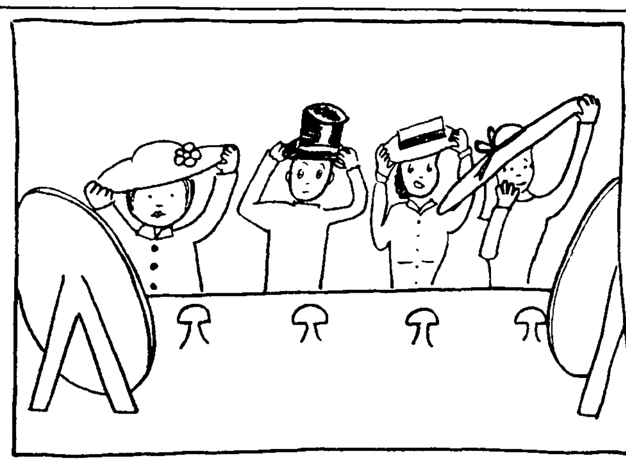
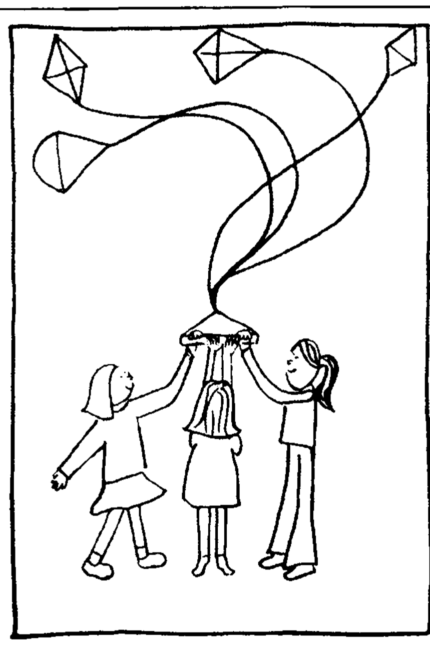
だれかがどの料理も食べてみた。
Dareka-ga dono ryouri-mo tabete-mita.
someone-NOM every dish-QPt tried
'Someone tried every dish.'</td><td>31</td><td>7</td></tr>
<tr><td>2b</td><td>

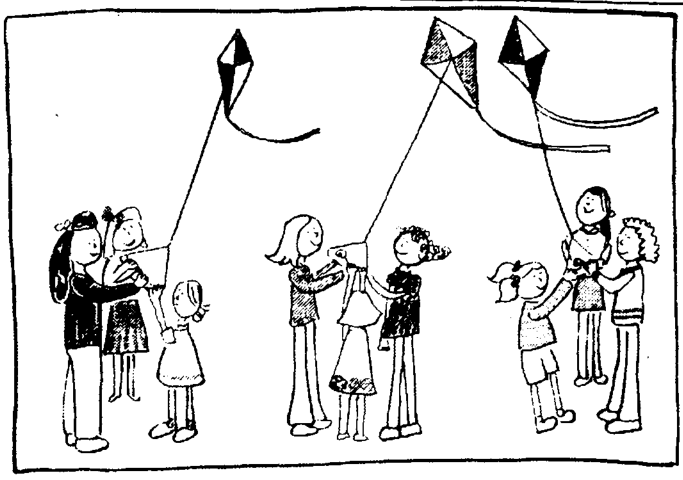
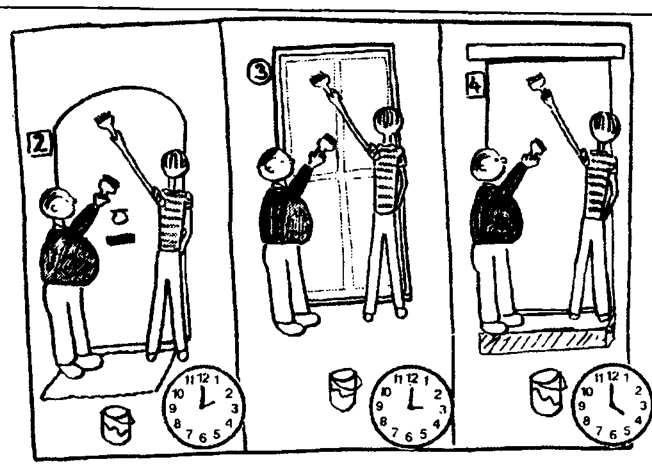
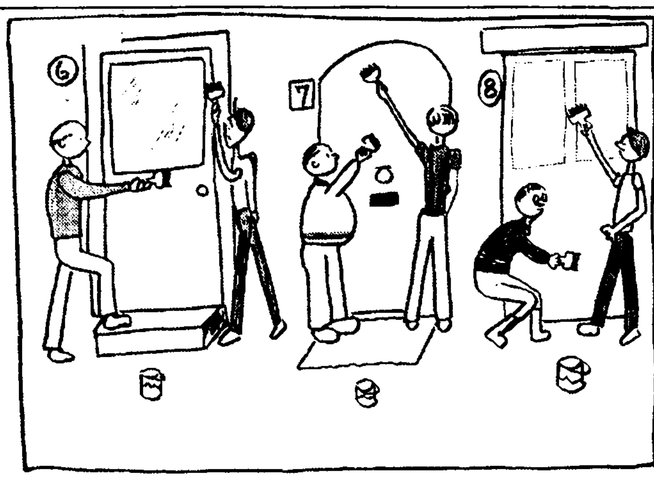
(Sentence as 2a)</td><td>27</td><td>15</td></tr>
<tr><td>3</td><td>
だれかがどの猫もなでた。
Dareka-ga dono neko-mo nadeta.
someone-NOM every cat-QPt stroked
'Someone stroked every cat.'</td><td>36</td><td>19</td></tr>
</table>

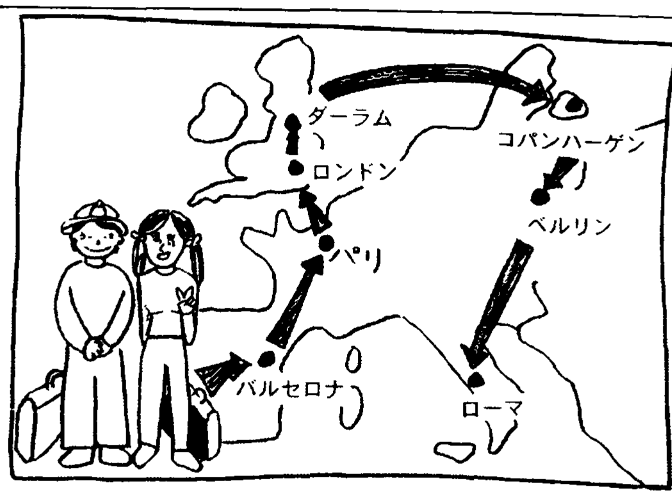
3b	 (Sentence as 3a)	19	5
4a	 だれかがどの傘[かさ]もさした。 Dareka-ga dono kasa-mo sasita. someone-NOM every umbrella-QPt put up 'Someone used every umbrella.'	22	43
4b	 (*Sentence as 4a*)	13	27

5a	 だれかがすべての本を読んだ。 Dareka-ga subete-no hon-o yonda. someone-NOM all-GEN book-ACC read 'Someone read all the books.'	9	17
5b	 (*Sentence as 5a*)	35	21
6a	 だれかがすべてのスーツケースを運んだ。 Dareka-ga subete-no suutukeesu-o hakonda. someone-NOM all-GEN suitcase-ACC carried 'Someone carried all the suitcases.'	26	13

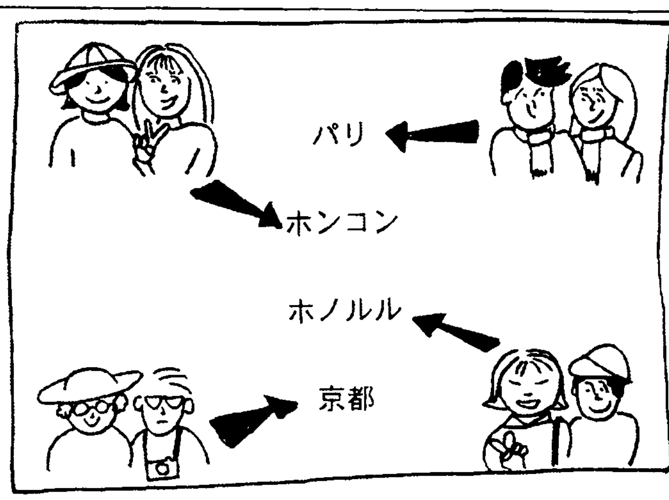
6b	 (Sentence as 6a)	32	34
7a	 だれかがすべてのプレゼントを開(あ)けた。 Dareka-ga subete-no purezento-o aketa. someone-NOM all-GEN present-ACC opened 'Someone opened all the presents.'	14	26
7b	 (Sentence as 7a)	11	6

8a	 だれかがすべての帽子を被ってみた。 Dareka-ga subete-no bousi-o kabutte-mita. someone-NOM all-GEN bousi-ACC tried on 'Someone tried on all the hats.'	12	37
8b	 (Sentence as 8a)	23	12
9a	 三人の女の子がどのたこも上げた。 Sannin-no onnanoko-ga dono tako-mo ageta. three.HUMAN-GEN girl-NOM every kite-QPt flew 'Three girls flew every kite.'	16	44

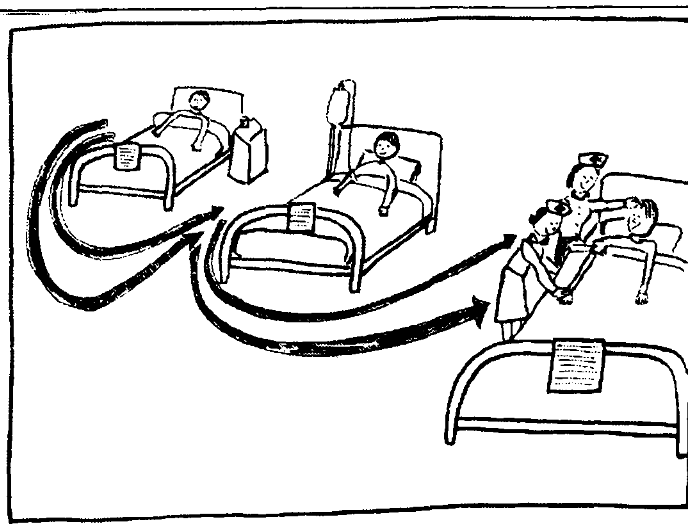
9b	 (*Sentence as 9a*)	7	20
10a	 二人の男がどのドアもぬった。 Hutari-no otoko-ga dono doa-mo nutta. two.HUMAN-GEN man-NOM every door-QPt painted Two men painted every door.	17	30
10b	 (*Sentence as 10a*)	6	8

<table>
<tr><td>11a</td><td>

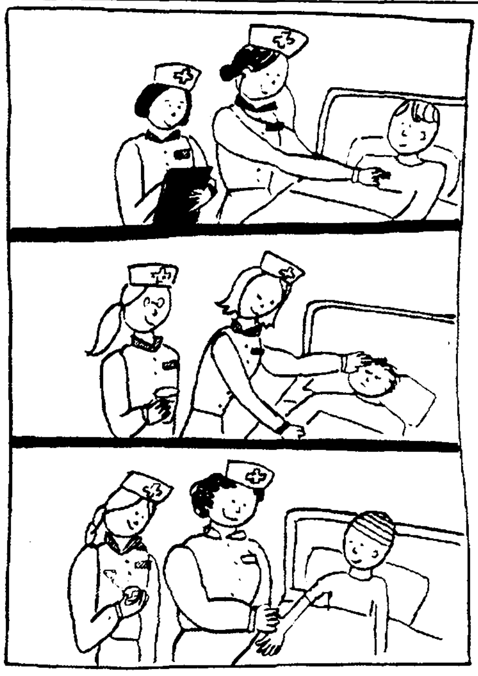
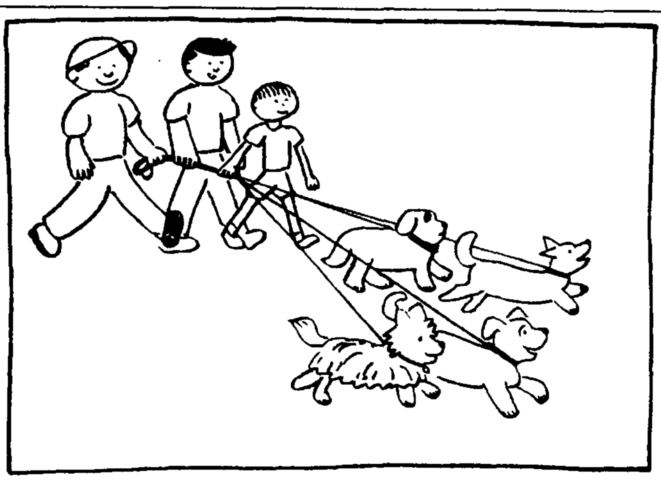
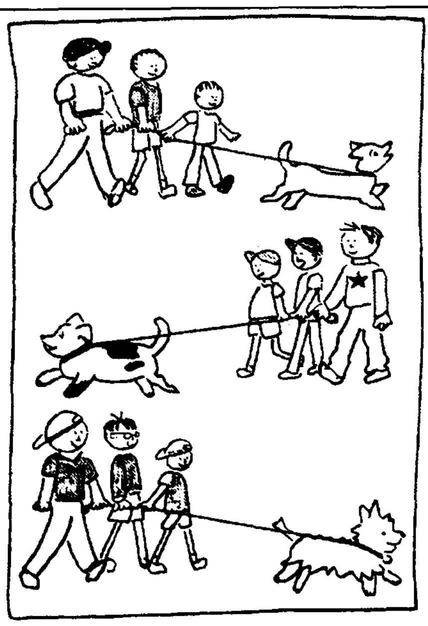
Translation (following arrows): Barcelona, Paris, London, Durham, Copenhagen, Berlin, Rome

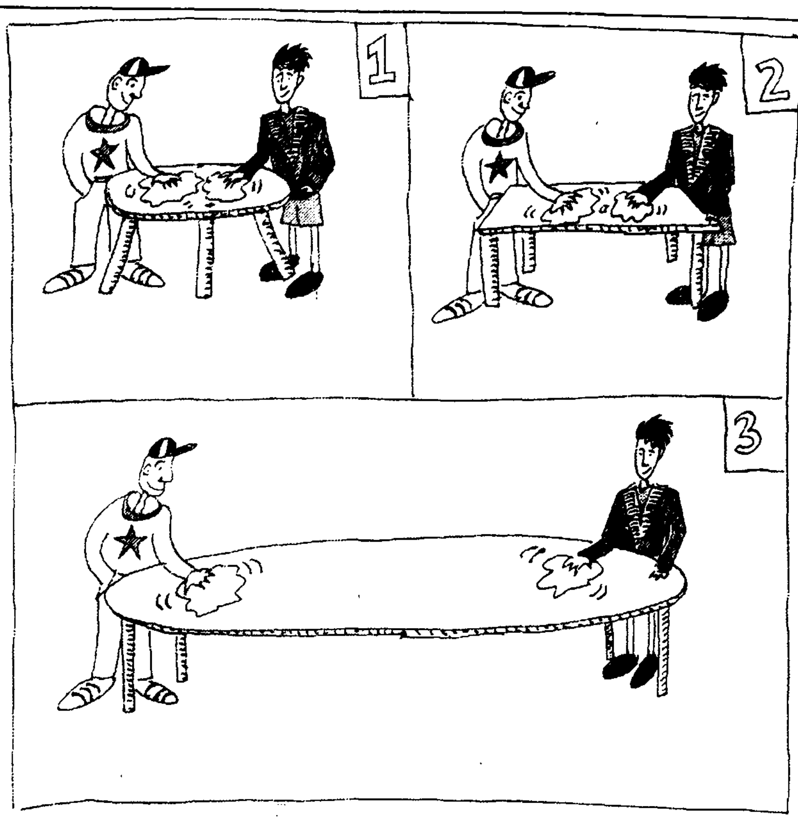
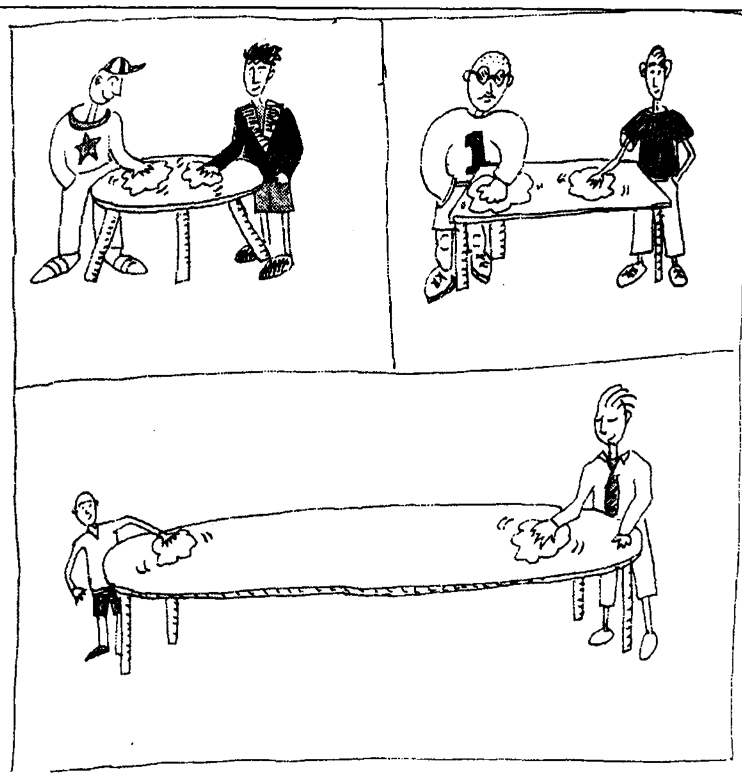
二人の観光客がどの町も見物した。
Hutari-no kankoukyaku-ga dono mati-mo kenbutu-sita.
two.HUMAN-GEN tourist-NOM every town-QPt visited
'Two tourists visited every city.'</td><td>28</td><td>11</td></tr>
<tr><td>11b</td><td>

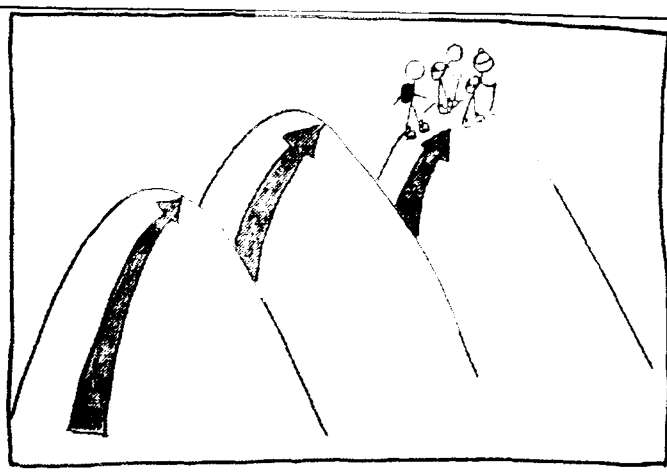
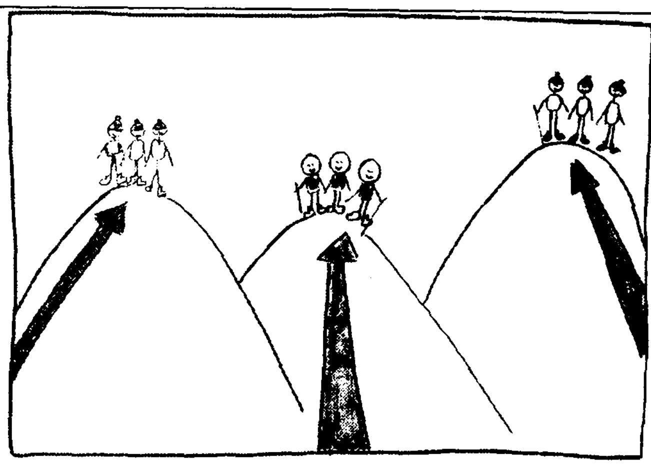
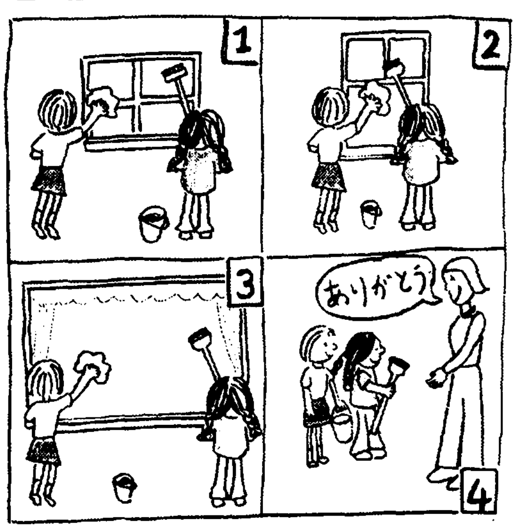
Translation: Paris, Hong Kong, Honolulu, Kyoto

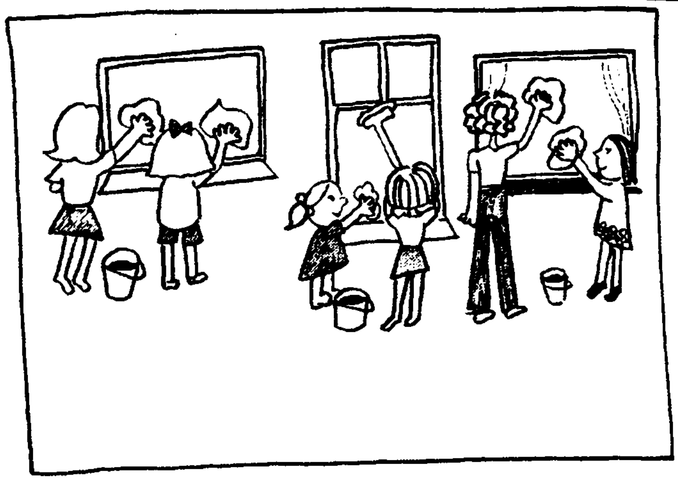
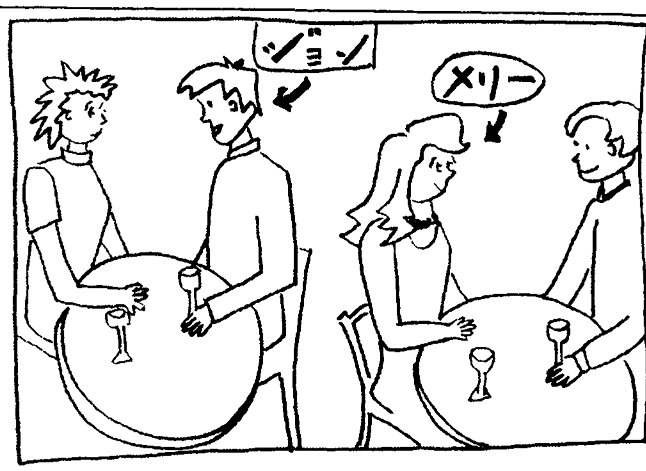
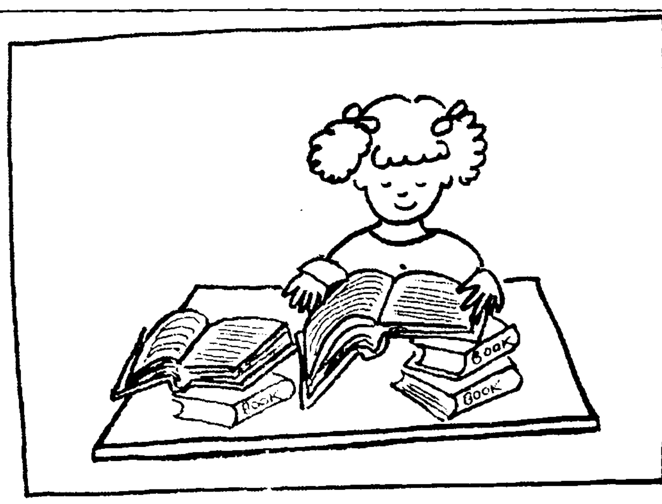
(Sentence as 11a)</td><td>10</td><td>36</td></tr>
<tr><td>12a</td><td>

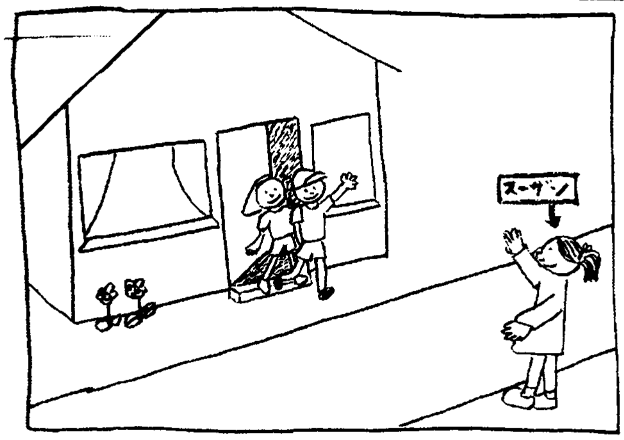
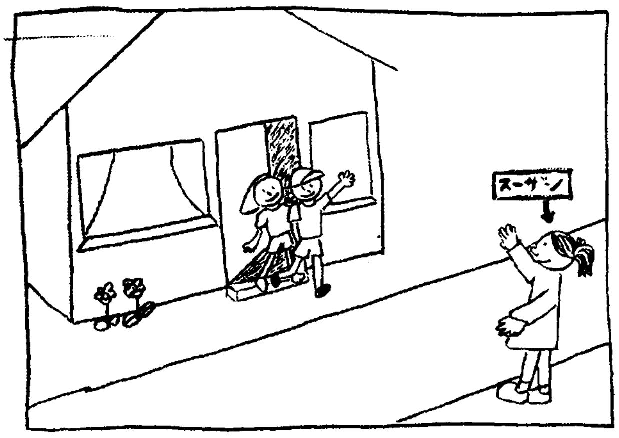
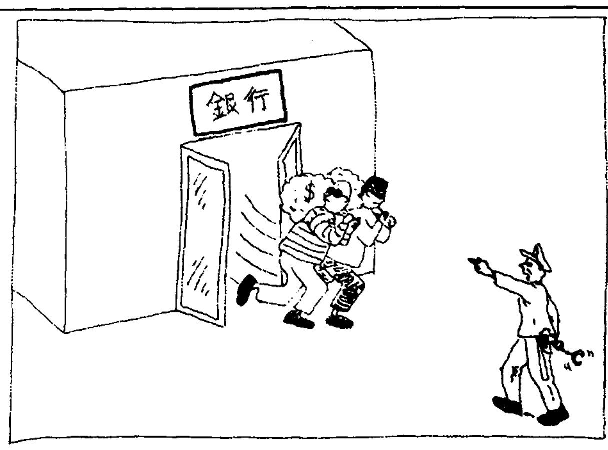
二人の看護婦がどの患者も看病した。
Hutari-no kangohu-ga dono kanzya-mo kanbyou-sita.
two.HUMAN-GEN nurse-NOM every patient-QPt looked after
'Two nurses looked after every patient.'</td><td>39</td><td>40</td></tr>
</table>

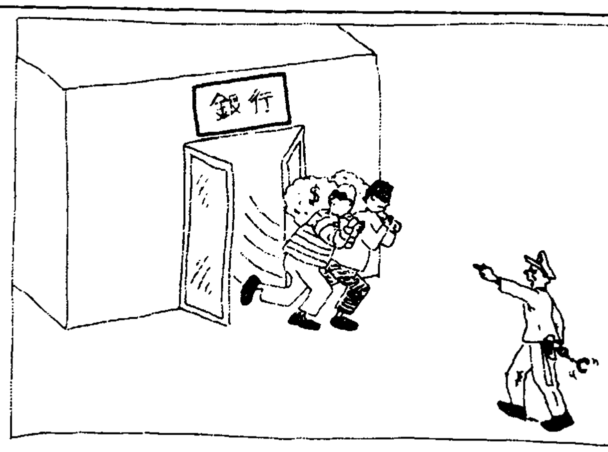
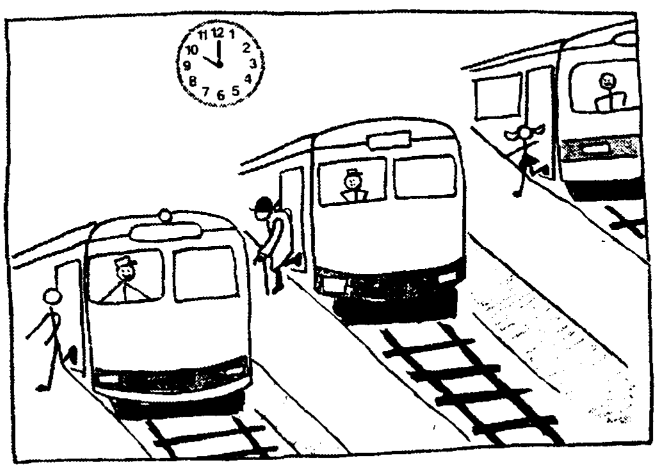
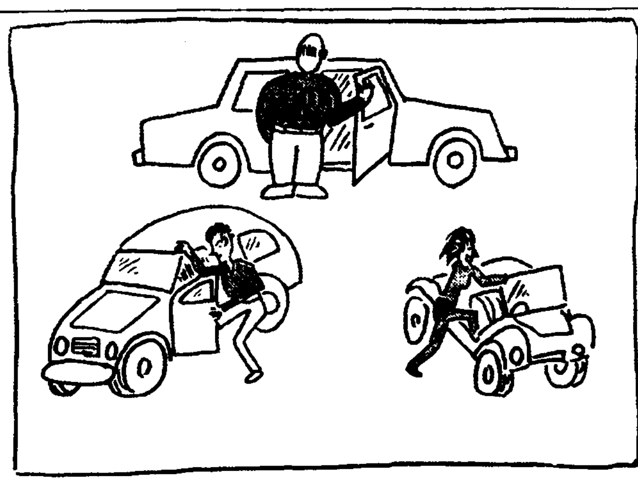
12b	 (Sentence as 12a)	20	22
13a	 三人の男の子がすべての犬を散歩に連れて行った Sannin-no otokonoko-ga subete-no inu-o sanpo ni turete-itta. three.HUMAN-GEN boy-NOM all-GEN dog-ACC walk for took 'Three boys walked all the dogs.'	8	23
13b	 (Sentence as 13a)	44	31

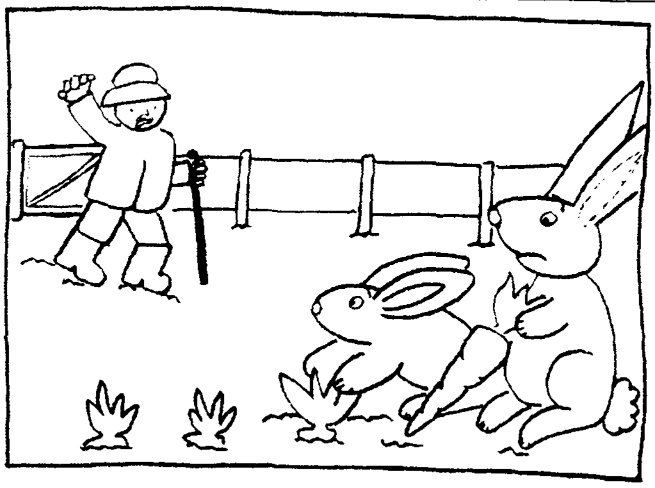
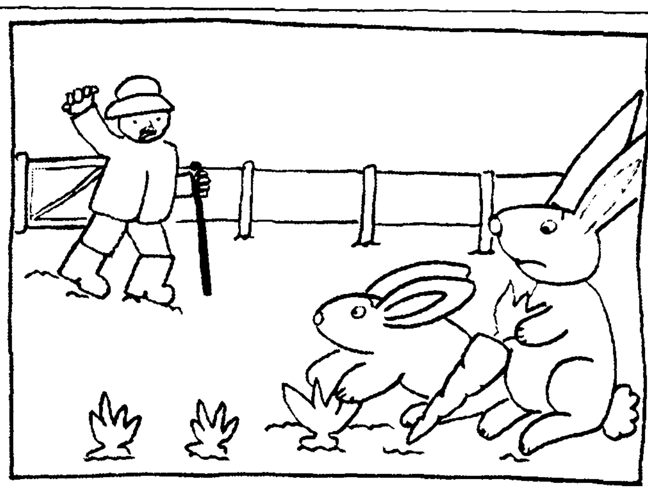
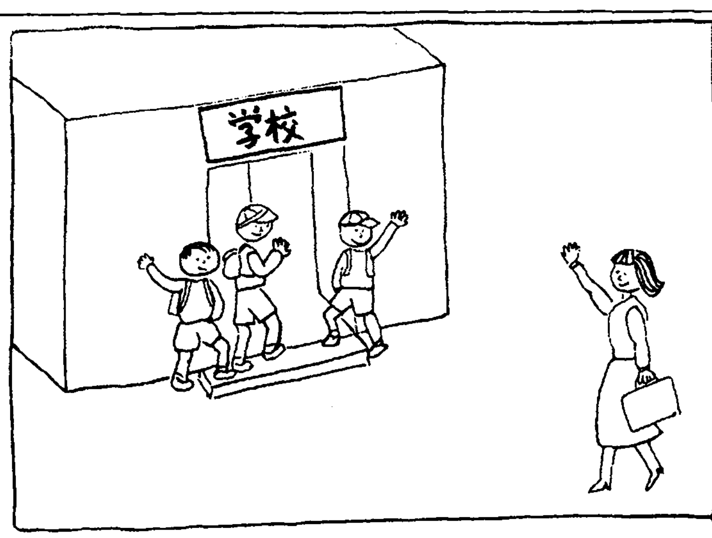
14a	 二人の男の子がすべてのテーブルを拭いた。 Hutari-no otokonoko-ga subete-no teeburu-o huita. two.HUMAN-GEN boy-NOM all-GEN table-ACC wiped Two boys wiped all the tables.	33	29
14b	 (Sentence as 14a)	4	14

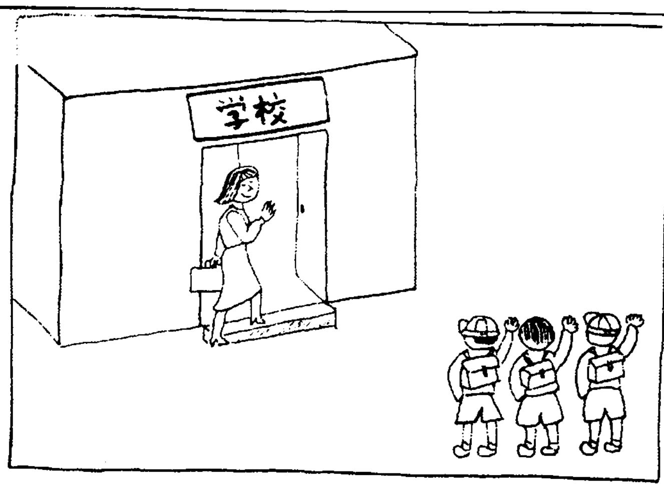
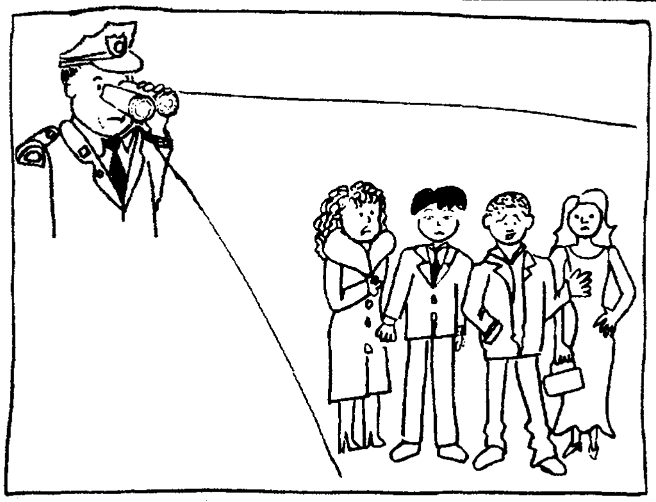
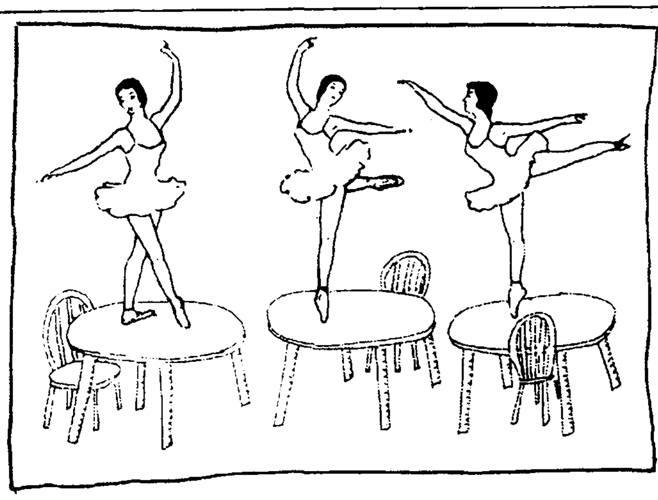
15a	 三人の登山客がすべての山に登った。 Sannin-no tozankyaku-ga subete-no yama-ni nobotta. three.HUMAN-GEN hiker-NOM all-GEN mountain-LOC climb 'Three hikers climbed all the mountains.'	5	16
15b	 (Sentence as 15a)	37	33
16a	 **Translation:** 'Thank you!' 二人の女の子がすべての窓を洗った。 Hutari-no onnanoko-ga subete-no mado-o aratta. two.HUMAN-GEN girl-NOM all-GEN window-ACC washed 'Two girls washed all the windows.'	42	41

16b	 (Sentence as 16a)	30	35
D01	 **Translation** (L-R): John, Mary ジョンとメリーがデートをした。 Zyon to Merii-ga deeto-o sita. John and Mary-NOM date-ACC had 'John and Mary went on a date.'	2	2
D02	 女の子が新聞(しんぶん)を読(よ)んだ。 Onnanoko-ga sinbun-o yonda. girl-NOM newspaper-ACC read 'A girl read the newspaper.'	1	1

D03	 Translation: Susan 子供が二人ともスーザンを見た。 Kodomo-ga hutari tomo Suuzan-o mita. Child-NOM two.HUMAN both Susan-ACC saw 'Both children saw Susan.'	21	24
D04	 Translation: Susan 子供が二人ともスーザンに手を振った。 Kodomo-ga hutari tomo Suuzan ni te-o hutta. Child-NOM two.HUMAN both Susan at hand-ACC waved 'Both children waved at Susan.'	34	39
D05	 **Translation** (above door)**:** Bank 警察官が二人の泥棒を見た。 Keisatukan-ga hutari-no dorobou-o mita. police officer-NOM two.HUMAN-GEN thief-ACC saw 'A policeman saw two thieves.'	3	25

D06	 **Translation** (above door): Bank 警察官が二人とも泥棒を見た。 Keisatukan-ga hutari tomo dorobou-o mita. police officer-NOM two.HUMAN both thief-ACC saw 'Both policemen saw the thieves.'	25	42
D07	 三人が同じ電車を降りた。 Sannin-ga onazyi densya-o orita. three people-NOM same train-ACC got off 'Three people got off the same train.'	18	9
D08	 皆が同じ車に乗った。 Minna-ga onazyi kuruma ni notta. everyone-NOM same car in got 'Everyone got in the same car.'	40	32

D09	 だれかが二匹のウサギを見た。 Dareka-ga ni-hiki-no usagi-o mita. someone-NOM two-CL-GEN rabbit-ACC saw 'Someone saw the two rabbits.'	15	38
D10	 ウサギが二匹ともにんじんを盗んだ。 Usagi-ga ni-hiki tomo ninzyin-o nusunda. rabbit-NOM two-CL both carrot-ACC stole 'Both rabbits stole the carrots.'	43	3
D11	 **Translation** (above door): School 三人の子供が先生に手を振った。 Sannin-no kodomo-ga sensei ni te-o hutta. three.HUMAN-GEN child-NOM teacher at hand-ACC waved 'Three children waved at the teacher.'	38	28

D12	 **Translation** (above door): School 先生が二人の子供に手を振った。 Sensei-ga hutari-no kodomo ni te-o hutta. teacher-NOM two.HUMAN-GEN child at hand-ACC waved 'The teacher waved at both children.'	24	10
D13	 警察官が五人を見た。 Keisatukan-ga gonin-o mita. police officer-NOM five people-ACC watched 'The detective watched five people.'	45	45
D14	 三人のバレリーナが椅子の上で踊った。 Sannin-no bareriina-ga isu-no ue de odotta. three.HUMAN-GEN ballerina-NOM chair-GEN top on danced 'Three ballerinas danced on the chairs.'	46	46

Appendix 2B: Answer sheet for learners of Japanese, Phase 2A

Personal Details

1. Your initials:................................ 2. Your age: ....
3. What is (are) your native language(s)? ....

   If you have more than one native language, which language is dominant, if any? Please give details:..

   ....

4. Please describe your formal study of Japanese by filling in this table:

Type of instruction	Duration
Example: *degree in Japanese at UK university, approx. 8 hrs per week of instruction*	*2 years 5 months*

5. Have you ever lived in Japan? yes no

   If yes, how long did you live there? ....

****

Instructions

This test involves judging how well particular sentences match particular pictures. Each picture will be shown on the overhead projector and you will have a few seconds to look at it. A sentence will then be shown with the picture. Please decide how well the sentence matches the picture, and indicate your answer by circling a number on the scale in your answersheet. The scale ranges from 'no, (it does) not (match) at all' to 'yes, (it matches) perfectly', as follows:

(No, not at all) −2 −1 0 +1 +2 (Yes, perfectly)

(Note: all the sentences in the test are grammatically correct.)

1

Examples

(Does the sentence match the picture?)

	No, not at all				Yes, perfectly
Ex. 1	-2	-1	0	+1	+2
Ex. 2	-2	-1	0	+1	+2
Ex. 3	-2	-1	0	+1	+2
Ex. 4	-2	-1	0	+1	+2
Ex. 5	-2	-1	0	+1	+2
Ex. 6	-2	-1	0	+1	+2

****

Your comments (after completing the actual task on pages 3 & 4)

(Please circle the relevant option, where options are given)

1. Please comment on the time allowed for viewing the picture before seeing the sentence:

   too long just right too short

2. Please comment on the time allowed for viewing the picture and sentence together:

   too long just right too short

3. Please comment on the total length of the test: too long OK
4. Did you find any of the pictures difficult to understand?

   ....

   ....

5. Any further comments?

   ....

   ....

   ....

2

(Does the sentence match the picture?)

	No, not at all				Yes, perfectly
1	-2	-1	0	+1	+2
2	-2	-1	0	+1	+2
3	-2	-1	0	+1	+2
4	-2	-1	0	+1	+2
5	-2	-1	0	+1	+2
6	-2	-1	0	+1	+2
7	-2	-1	0	+1	+2
8	-2	-1	0	+1	+2
9	-2	-1	0	+1	+2
10	-2	-1	0	+1	+2
11	-2	-1	0	+1	+2
12	-2	-1	0	+1	+2
13	-2	-1	0	+1	+2
14	-2	-1	0	+1	+2
15	-2	-1	0	+1	+2
16	-2	-1	0	+1	+2
17	-2	-1	0	+1	+2
18	-2	-1	0	+1	+2
19	-2	-1	0	+1	+2
20	-2	-1	0	+1	+2
21	-2	-1	0	+1	+2
22	-2	-1	0	+1	+2
23	-2	-1	0	+1	+2
24	-2	-1	0	+1	+2
25	-2	-1	0	+1	+2

3

(Does the sentence match the picture?)

	No, not at all				Yes, perfectly
26	-2	-1	0	+1	+2
27	-2	-1	0	+1	+2
28	-2	-1	0	+1	+2
29	-2	-1	0	+1	+2
30	-2	-1	0	+1	+2
31	-2	-1	0	+1	+2
32	-2	-1	0	+1	+2
33	-2	-1	0	+1	+2
34	-2	-1	0	+1	+2
35	-2	-1	0	+1	+2
36	-2	-1	0	+1	+2
37	-2	-1	0	+1	+2
38	-2	-1	0	+1	+2
39	-2	-1	0	+1	+2
40	-2	-1	0	+1	+2
41	-2	-1	0	+1	+2
42	-2	-1	0	+1	+2
43	-2	-1	0	+1	+2
44	-2	-1	0	+1	+2
45	-2	-1	0	+1	+2
46	-2	-1	0	+1	+2

ありがとう THANK YOU! ありがとう

4

Appendix 2C: Tables of results for Phase 2A

Table 2C.1: Mean ratings on test types by participant

	id	age	order	mean rating (scale: –2 to 2)							
				1a	1b	2a	2b	3a	3b	4a	4b
Native Japanese	JJ01	28	1	1.75	0	1.75	0.25	2	0.5	2	1
	JJ02	29	1	2	-1.75	1.75	-1.5	2	-1.25	2	-1.25
	JJ03	29	1	1	-1	1	-1.75	1.75	0.25	1.75	-1.25
	JJ04	19	2	1.75	-0.75	2	-1	2	-0.25	2	-1
	JJ05	18	2	0.75	-1.67	0.75	-1.5	0.75	-1.75	1	-2
	JJ06	18	2	1	-2	2	-2	2	-2	2	-2
	JJ08	19	2	1.5	-1.5	2	-1.75	1.5	-1	1.75	-1.25
	JJ09	18	2	1.75	0.75	2	1	2	1.25	2	2
	JJ10	18	2	1.75	-0.25	2	-0.25	2	-1.75	2	-2
	JJ11	18	2	1.5	-0.5	1	0	0.75	-0.25	2	-1
	JJ12	18	2	2	-1.25	1.75	-1.25	0.75	-1	1.75	-0.75
	JJ13	19	2	1.25	0.25	1.667	-0.75	1.25	-1	2	-0.5
	JJ14	18	2	1.75	-1.5	1.75	-1.5	1.5	-1.5	2	-1.5
	JJ15	18	2	1.25	-0.5	1.5	0.25	1	-0.5	1.5	-1
	JJ17	18	2	2	-2	2	-2	2	-2	2	-2
	JJ18	18	2	1.75	-1	2	-0.25	1.75	-2	2	-1.75
	JJ21	18	2	1.75	-1.75	2	-1.75	2	-2	2	-2
	JJ22	19	2	1.25	-0.5	2	-0.5	2	0	2	-0.25
	JJ24	21	2	1.5	-1	2	-1	1.25	-1.25	2	-1.25
	JJ27	18	2	1.75	-1	2	-2	1	-2	2	-2
	JJ28	18	2	1	-1	2	-2	1.25	-2	2	-2
	JJ29	19	2	1.5	0	2	-0.75	2	0	2	-1.5
	JJ30	18	2	1.5	-1.25	2	-1.75	1.75	-2	2	-2
	JJ31	18	2	1	-2	2	-2	1	-2	2	-2
	JJ33	20	2	1.75	-1	2	-0.25	2	-1.75	2	-2
L2 Japanese	EJ01	29	1	1.75	-1.5	1.75	-1.25	1	0.75	2	-0.25
	EJ02	22	1	2	-2	2	-2	2	-2	2	-2
	EJ03	22	1	2	-1	2	-0.25	2	1	2	-1.5
	EJ04	25	1	2	-2	2	-1.75	2	-1.25	2	-1.25
	EJ05	22	1	2	-2	1.75	-2	2	1.25	2	-1
	EJ06	21	1	2	1.5	2	1.25	2	1	2	1
	EJ07	21	1	2	-2	0.5	-2	2	-2	2	-2
	EJ08	22	1	2	0.25	2	-1.5	1.75	-2	1.25	-0.5
	EJ09	26	1	2	-1.5	2	-1.25	2	0	1.75	-1.25
	EJ10	25	1	1	0.25	2	-1.5	1.25	1.5	2	0.25
	EJ11	22	1	1.25	-1.25	2	-1.5	2	-1	2	-0.75
	EJ12	22	1	2	1	2	0	2	1	1.75	0
	EJ13	23	1	1.5	1	2	-0.5	1	0.5	2	-1.25
	EJ14	22	1	2	-1	2	0.5	1	-2	2	-0.75
	EJ15	34	2	1.25	0	1	-0.5	2	0	2	-1.75
	EJ16	22	2	1.75	2	2	0	2	-0.5	2	-1
	EJ17	23	2	1.5	0.75	2	-0.5	1.5	0	2	-1.5
	EJ18	29	2	2	-0.5	2	-1.25	2	-1.5	2	-1.75

	id	age	order	mean rating (scale: –2 to 2)							
				1a	1b	2a	2b	3a	3b	4a	4b
Native English	EE01	23	1	1.25	2	1.75	0.5	2	1.75	2	0.5
	EE02	21	1	1	0.75	1	-2	2	-1.25	2	-2
	EE03	29	1	2	-2	2	-2	2	-2	2	-2
	EE04	29	1	2	1	2	1	2	1.5	2	-1.25
	EE05	29	1	0.5	1	1.5	-1	1	2	2	-1.25
	EE06	37	1	2	1.25	2	-0.5	2	1	2	0.5
	EE07	32	1	1.25	0.5	2	-1	2	-1	2	-1.75
	EE08	20	1	2	-2	2	-2	2	-2	2	-2
	EE09	21	1	1.25	0.5	2	-1	1.75	1.75	2	-1
	EE10	48	1	2	1	1	-1	2	2	1	1
	EE11	51	1	1.75	2	2	1	2	1.75	1.75	0.5
	EE12	26	1	2	2	1.75	2	2	1.75	2	1.25
	EE13	44	1	1.75	-1.75	1.75	-1.75	2	-1.25	1.5	-1.75
	EE14	41	1	1.25	0.25	1.5	-1	1.25	-0.5	1.5	-1
	EE15	35	1	2	-2	2	-2	2	-0.5	2	-1.75
	EE16	20	1	2	-0.25	2	-2	1.5	2	2	-1.25
	EE17	26	2	0.75	0.25	2	-2	2	0	2	-0.25
	EE18	18	2	1.25	1	2	-2	1.75	1	2	-1
	EE19	20	2	1.25	1	1	0.5	1.75	-0.25	2	-0.5
	EE20	18	2	0.25	1.75	2	-0.75	1.5	-1	2	-1
	EE21	19	2	2	0.75	2	-2	1.5	0.25	2	-1.25
	EE22	20	2	1.75	1.5	2	-1	1.75	0.25	2	-0.25
	EE23	21	2	1.5	1	1.5	-0.75	1.5	0	2	-1.75
	EE24	19	2	1.75	2	1.75	1.75	2	2	2	-1
	EE25	22	2	2	0.5	2	-0.75	2	0	2	-1
	EE26	20	2	0.75	1.75	2	-1.75	2	-1.75	2	-1.75
	EE27	18	2	1	1	2	-1.25	1.5	1.25	2	-0.5
	EE28	18	2	1.75	1.75	2	-0.5	1	1.75	2	1.5
	EE29	19	2	1.5	0.5	2	-1.25	1.75	0.25	2	-1.75

Table 2C.2: Mean ratings on test tokens by participant group

index	JJ		EJ		EE	
	mean	SD	mean	SD	mean	SD
1A	0.63	1.26	1.56	0.86	1.28	0.88
1B	-1	1.2	-0.3	1.64	0.9	1.4
2A	1.84	0.37	1.94	0.24	1.72	0.53
2B	-1.4	1.12	-0.6	1.58	0.86	1.41
3A	1.68	0.82	1.83	0.51	1.55	0.69
3B	-0.6	1.46	-0.1	1.92	0.62	1.4
4A	1.95	0.23	1.78	0.43	1.45	0.83
4B	-0.7	1.41	-0.8	1.63	0.24	1.5
5A	1.78	0.43	1.89	0.32	1.83	0.38
5B	-1.3	1.15	-1.3	1.27	-1	1.38
6A	1.79	0.71	2	0	1.9	0.31
6B	-1.3	1.05	-0.9	1.26	-1	1.38
7A	2	0	1.94	0.24	1.76	0.51
7B	-0.4	1.57	-0.9	1.39	-0.7	1.36
8A	1.95	0.23	1.5	1.29	1.76	0.79
8B	-1	1.11	-0.4	1.38	-0.6	1.48
9A	1.58	0.96	1.44	1.2	1.52	0.69
9B	-1	1.45	0.12	1.69	0.14	1.6
10A	2	0	2	0	1.93	0.26
10B	-1.3	1.15	0	1.78	0.31	1.63
11A	0.58	1.61	1.61	1.14	1.72	0.7
11B	-1.6	0.6	-0.9	1.37	0.1	1.68
12A	2	0	1.94	0.24	1.93	0.26
12B	-1.1	1.45	-0.4	1.61	0.93	1.46
13A	1.95	0.23	1.78	0.73	1.97	0.19
13B	-1.2	1.21	-1.2	1.17	-1	1.32
14A	2	0	2	0	2	0
14B	-1.4	1.12	-0.1	1.78	-0.6	1.5
15A	1.89	0.32	1.94	0.24	1.9	0.41
15B	-1.4	1.12	-1.4	1.14	-0.9	1.16
16A	2	0	2	0	1.83	0.76
16B	-1.4	0.96	-1.1	0.96	-0.8	1.42
D01	-1.8	0.41	-1.9	0.47	-1.6	1.09
D02	1.48	0.71	1.56	0.78	1.59	0.57
D03	1.88	0.33	2	0	1.9	0.41
D04	-0.3	1.74	-0.6	1.61	0.04	1.86
D05	1.88	0.33	1.94	0.24	1.97	0.19
D06	-1.1	1.54	-0.2	1.86	-1.9	0.37
D07	-1.9	0.33	-2	0	-1.9	0.74
D08	-2	0	-2	0	-2	0
D09	1.8	0.41	1.94	0.24	1.9	0.41
D10	-1.8	0.52	-1.7	0.67	-0.9	1.52
D11	1.96	0.2	2	0	1.97	0.19
D12	-1.9	0.28	-1.3	1.32	-1.9	0.41
D13	-2	0.2	-1.6	1.15	-1.7	0.88
D14	-1.8	0.8	-1.7	0.96	-1.8	0.77

2C.2. Repeated measures ANOVA[212]

Table 2C.3.i: Descriptive statistics

	GROUP	Mean	Std. Deviation	N
Type 1a	JJ	1.5100	.35707	25
	EJ	1.7778	.33087	18
	EE	1.5000	.51322	29
Type 1b	JJ	-.9667	.74068	25
	EJ	-.4444	1.32164	18
	EE	.6552	1.20899	29
Type 2a	JJ	1.7967	.36171	25
	EJ	1.8333	.41124	18
	EE	1.8103	.32497	29
Type 2b	JJ	-1.0400	.87106	25
	EJ	-.8889	.93629	18
	EE	-.8448	1.16958	29
Type 3a	JJ	1.5700	.47610	25
	EJ	1.7500	.40220	18
	EE	1.7759	.30869	29
Type 3b	JJ	-1.0900	.95982	25
	EJ	-.2917	1.26389	18
	EE	.3707	1.33734	29
Type 4a	JJ	1.9100	.22684	25
	EJ	1.9306	.18798	18
	EE	1.9224	.22264	29
Type 4b	JJ	-1.2500	.99216	25
	EJ	-.9583	.81462	18
	EE	-.8190	1.02403	29

[212] Sphericity is assumed, as Mauchly's test of sphericity was non-significant. Equality of variance is not assumed, as Levene's test included significant differences:

Levene's Test of Equality of Error Variances:

	F	df1	df2	Sig.
Type 1a	3.832	2	69	.026
Type 1b	4.025	2	69	.022
Type 2a	.018	2	69	.982
Type 2b	.413	2	69	.663
Type 3a	5.333	2	69	.007
Type 3b	2.731	2	69	.072
Type 4a	.149	2	69	.862
Type 4b	.641	2	69	.530

Table 2C.3.ii: Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Subj QP	Obj QP	Scope	Type III Sum of Squares	df	Mean Square	F	Sig.
Subj QP	Linear			.014	1	.014	.040	.841
Subj QP * Group	Linear			.182	2	.091	.257	.774
Error(Subj QP)	Linear			24.423	69	.354		
Obj QP		Linear		7.083	1	7.083	35.908	.000
Obj QP * Group		Linear		11.608	2	5.804	29.425	.000
Error(Obj QP)		Linear		13.610	69	.197		
Scope			Linear	789.289	1	789.289	473.960	.000
Scope * Group			Linear	21.040	2	10.520	6.317	.003
Error(Scope)			Linear	114.906	69	1.665		
Subj QP * Obj QP	Linear	Linear		.000	1	.000	.001	.971
Subj QP * Obj QP * Group	Linear	Linear		.098	2	.049	.263	.770
Error(Subj QP*Obj QP)	Linear	Linear		12.837	69	.186		
Subj QP * Scope	Linear		Linear	1.248	1	1.248	2.837	.097
Subj QP * Scope * Group	Linear		Linear	.632	2	.316	.718	.491
Error(Subj QP*Scope)	Linear		Linear	30.350	69	.440		
Obj QP * Scope		Linear	Linear	27.552	1	27.552	79.324	.000
Obj QP * Scope * Group		Linear	Linear	9.700	2	4.850	13.964	.000
Error(Obj QP*Scope)		Linear	Linear	23.966	69	.347		
Subj QP * Obj QP * Scope	Linear	Linear	Linear	.000	1	.000	.001	.981
Subj QP * Obj QP * Scope * Group	Linear	Linear	Linear	1.120	2	.560	2.043	.137
Error(Subj QP*Obj QP*Scope)	Linear	Linear	Linear	18.906	69	.274		

Table 2C.3.iii: Tests of Between-Subjects Contrasts

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	175.683	1	175.683	105.258	.000
Group	25.971	2	12.986	7.780	.001
Error	115.166	69	1.669		

Table 2C.4.i:
Between-groups post hoc multiple comparisons (Games-Howell): Type 1a

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ	-.2678(*)	.10575	.040	-.5256	-.0100
	EE	.0100	.11909	.996	-.2777	.2977
EJ	JJ	.2678(*)	.10575	.040	.0100	.5256
	EE	.2778	.12314	.073	-.0207	.5763
EE	JJ	-.0100	.11909	.996	-.2977	.2777
	EJ	-.2778	.12314	.073	-.5763	.0207

Table 2C.4.ii:
Between-groups post hoc multiple comparisons (Games-Howell): Type 1b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ	-.5222	.34494	.302	-1.3821	.3377
	EE	-1.6218(*)	.26897	.000	-2.2727	-.9710
EJ	JJ	.5222	.34494	.302	-.3377	1.3821
	EE	-1.0996(*)	.38398	.019	-2.0409	-.1583
EE	JJ	1.6218(*)	.26897	.000	.9710	2.2727
	EJ	1.0996(*)	.38398	.019	.1583	2.0409

Table 2C.4.iii:
Between-groups post hoc multiple comparisons (Games-Howell): Type 2a

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ	-.0367	.12095	.951	-.3331	.2598
	EE	-.0137	.09421	.988	-.2414	.2140
EJ	JJ	.0367	.12095	.951	-.2598	.3331
	EE	.0230	.11418	.978	-.2585	.3045
EE	JJ	.0137	.09421	.988	-.2140	.2414
	EJ	-.0230	.11418	.978	-.3045	.2585

Table 2C.4.iv:
Between-groups post hoc multiple comparisons (Games-Howell): Type 2b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ	-.1511	.28116	.853	-.8391	.5369
	EE	-.1952	.27842	.764	-.8673	.4769
EJ	JJ	.1511	.28116	.853	-.5369	.8391
	EE	-.0441	.30963	.989	-.7963	.7082
EE	JJ	.1952	.27842	.764	-.4769	.8673
	EJ	.0441	.30963	.989	-.7082	.7963

Table 2C.4.v:
Between-groups post hoc multiple comparisons (Games-Howell): Type 3a

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ	-.1800	.13436	.382	-.5071	.1471
	EE	-.2059	.11114	.166	-.4764	.0646
EJ	JJ	.1800	.13436	.382	-.1471	.5071
	EE	-.0259	.11078	.970	-.2993	.2476
EE	JJ	.2059	.11114	.166	-.0646	.4764
	EJ	.0259	.11078	.970	-.2476	.2993

Table 2C.4.vi:
Between-groups post hoc multiple comparisons (Games-Howell): Type 3b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ	-.7983	.35440	.078	-1.6715	.0748
	EE	-1.4607(*)	.31388	.000	-2.2186	-.7027
EJ	JJ	.7983	.35440	.078	-.0748	1.6715
	EE	-.6624	.38784	.216	-1.6085	.2837
EE	JJ	1.4607(*)	.31388	.000	.7027	2.2186
	EJ	.6624	.38784	.216	-.2837	1.6085

Table 2C.4.vii:
Between-groups post hoc multiple comparisons (Games-Howell): Type 4a

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ	-.0206	.06341	.944	-.1749	.1338
	EE	-.0124	.06138	.978	-.1606	.1358
EJ	JJ	.0206	.06341	.944	-.1338	.1749
	EE	.0081	.06060	.990	-.1393	.1555
EE	JJ	.0124	.06138	.978	-.1358	.1606
	EJ	-.0081	.06060	.990	-.1555	.1393

Table 2C.4.viii:
Between-groups post hoc multiple comparisons (Games-Howell): Type 4b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ	-.2917	.27612	.546	-.9636	.3802
	EE	-.4310	.27484	.268	-1.0944	.2323
EJ	JJ	.2917	.27612	.546	-.3802	.9636
	EE	-.1394	.27024	.864	-.7958	.5171
EE	JJ	.4310	.27484	.268	-.2323	1.0944
	EJ	.1394	.27024	.864	-.5171	.7958

Appendix 2D: Test sentences and written contexts for Phase 2B

Note: The table below presents the written Japanese context for each test item, along with its English translation, without romanised Japanese and glosses, since these are not necessary for understanding the test. Glosses of the test sentences are also omitted, as they are the same as for Phase 2A (see Appendix 2A).

index	context and test sentence	item no.
1a	昼休みに生徒が机に絵をかいているのを、先生が見つけた。その先生は生徒を一人ずつ叱った。 During school lunchtime, a teacher found a group of pupils drawing on the desks in a classroom. The teacher told each of the children off one by one. つまり、だれかがどの子供も叱った。 Tumari, dareka-ga dono kodomo-mo sikatta. 'In other words, someone scolded every child.'	21
1b	ひいおばあさんのお茶会で、三人の子供がいたずらをした。一人の大人が一人の男の子を叱り、もう一人の大人はもう一人の男の子を叱った。そして、三人目の大人が一人の女の子を叱った。 During Great Grandmother's tea party, three children were misbehaving. One adult told off one boy, another adult told off another boy, and a third adult told off a girl. *(Test sentence as 1a)*	12
1c	*(Context as 1a)* つまり、どの子供もだれかが叱った。 Tumari, dono kodomo-mo dareka-ga sikatta. 'In other words, someone scolded every child.' (*scrambled*)	5
1d	*(Context as 1b)* *(Test sentence as 1c)*	
2a	先週、ある男が街角にある喫茶店で、メニューにある料理を全部注文した。その男は一番おいしい料理がどれか知りたいのだと言って、全部の料理を食べてみた。 'Last week, a man ordered one of everything on the menu at the corner café. He said he wanted to see which dish was tastiest, and he tried each one.' つまり、だれかがどの料理も食べてみた。 Tumari, dareka-ga dono ryouri-mo tabete-mita. 'In other words, someone tried every dish.'	30
2b	『ネッドの食堂』には、十種類の料理がメニューにある。昨日の昼、十人の客が入ってきて、みんな違う料理を注文して、食べた。 'There are ten items on the menu at Ned's Diner. At lunchtime yesterday, ten customers came in and each chose a different item from the menu.' *(Test sentence as 2a)*	14

2c	*(Context as 2a)* つまり、どの料理もだれかが食べてみた。 Tumari, dono ryouri-mo dareka-ga tabeta. 'In other words, someone tried every dish. (*scrambled*)'	
2d	*(Context as 2b)* *(Test sentence as 2c)*	18
3a	朝日農場には猫が何匹かいる。昨日、この農場に小学生のグループが見学に来た。一人の小学生は、猫をみんな集めて、なでた。 'There are several cats on Asahi Farm. Yesterday, a group of primary school children went on a school trip to the farm. One of the children gathered all the cats together and stroked them.' つまり、だれかがどの猫もなでた。 Tumari, dareka-ga dono neko-mo nadeta. 'In other words, someone stroked every cat.'	42
3b	朝日農場には猫が何匹かいる。昨日、この農場に小学生のグループが見学に来た。一人の小学生は、猫をみんな集めて、なでた。 'Five cats were abandoned in a box outside the school gates. A group of five children found them. Each child picked up a different cat from the box and stroked it.' *(Test sentence as 3a)*	33
3c	*(Context as 3a)* つまり、どの猫もだれかがなでた。 Tumari, dono neko-mo dareka-ga nadeta. 'In other words, someone stroked every cat. (*scrambled*)'	36
3d	*(Context as 3b)* *(Test sentence as 3c)*	19
4a	昨日は雨が降っていた。男の子が一人、家中の傘を持ち出し、全部の傘をいっぺんにさしてみて、どのくらい濡れないでいるか試してみた。 'It was raining yesterday. A boy took all the umbrellas in the house and put them all up at once, to see how dry he could stay.' つまり、だれかがどの傘もさした。 Tumari, dono kasa-mo dareka-ga sasita. 'In other words, someone put up every umbrella.'	7
4b	レークビューホテルには客用の傘が十本ある。昨日は雨が降っていたので、十人の客は各自、一本ずつ傘を借りた。 'Lakeview Hotel has ten spare umbrellas for guests. Yesterday it was raining and ten different guests each borrowed an umbrella.' *(Test sentence as 4a)*	23

4c	*(Context as 4a)* つまり、どの傘もだれかがさした。 Tumari, dono kasa-mo dareka-ga sasita. 'In other words, someone put up every umbrella. (*scrambled*)'	34
4d	*(Context as 4b)* *(Test sentence as 4c)*	31
9a	さくら小学校には休み時間に子供たちが遊べるように、いくつか凧がある。昨日、三人の女の子が凧を全部出して、いっぺんにその凧を揚げた。 'Sakura Primary School has several kites for the pupils to play with during playtime. Yesterday, three girls took out all the kites and flew them together, all at once.' つまり、三人の女の子がどの凧も揚げた。 Tumari, sannin-no onnanoko-ga dono tako-mo ageta. 'In other words, three girls flew every kite.'	27
9b	富士見中学校は凧揚げ祭りをした。生徒たちは三つの大きい凧を使って、凧揚げのデモンストレーションをした。そのデモで、九人の女生徒が三人一組のチームになって、一つずつ凧を揚げた。 'Fujimi School held a kite festival. The pupils put on a kite-flying display with three big kites. Nine girls flew the kites in teams of three to one kite.' (Test sentence as 9a)	13
9c	*(Context as 9a)* つまり、どの凧も三人の女の子が揚げた。 Tumari, dono tako-mo sannin-no onnanoko-ga ageta. 'In other words, three girls flew every kite.' (*scrambled*)	16
9d	*(Context as 9b)* *(Test sentence as 9c)*	10
10a	先週、サクラ通りに面しているドアが全部塗りかえられた。二人の男がドアを一つずつ塗っていった。 'All the front doors in Sakura Street were painted last week. Two men went from one door to the next, painting each door.' つまり、二人の男がどのドアも塗った。 Tumari, hutari-no otoko-ga dono doa-mo nutta. 'In other words, two men painted every door.'	22

10b	村の公民館のドアを塗り直さなければならなかった。公民館にはドアが十個あったので、二十人の男性がボランティアをした。二人一組になって、それぞれドアを塗った。 'The doors of the village community centre needed to be repainted. The centre had 10 doors, and 20 men volunteered to paint them. They formed pairs and each pair painted one door.' *(Test sentence as 10a)*	28
10c	*(Context as 10a)* つまり、どのドアも二人の男が塗った。 Tumari, dono doa-mo hutari-no otoko-ga nutta. 'In other words, two men painted every door (*scrambled*).'	6
10d	*(Context as 10b)* *(Test sentence as 10c)*	20
11a	観光客の一団がツアーでフランスに行った。ツアーには六ヶ所の街へ行けるオプショナルツアーがあった。二人の観光客だけが、そのオプショナルツアーに参加して、六つの街を全部まわった。 'A group of tourists went on a package holiday to France. The package included optional visits to six different towns. Only two of the tourists took up the options and visited all six towns.' つまり、二人の観光客がどの街も見物した。 Tumari, hutari-no kankoukyaku-ga dono mati-mo kenbutu sita. 'In other words, two tourists visited every town.'	3
11b	八人の観光客がツアーでフィレンツェに行った。ツアーには四ヶ所の街へ行けるオプショナルツアーがあった。観光客はゆっくりしたかったので、各自、一ヶ所の街へ旅行した。二名はヴェニスに、二名はミラノに、二名はローマに、二名はピサに行った。 'Eight tourists went on a package holiday to Florence. The package included optional trips to four different towns. The tourists wanted to relax so they only took one trip each. Two of them went to Venice, two to Milan, two to Rome and two to Pisa.' *(Test sentence as 11a)*	29
11c	*(Context as 11a)* つまり、どの街も二人の観光客が見物した。 Tumari, dono mati-mo hutari-no kankoukyaku-ga kenbutu sita. 'In other words, two tourists visited every town (*scrambled*).'	15
11d	*(Context as 11b)* *(Test sentence as 11c)*	39

12a	先週、中央病院の看護婦が、二人以外はみんな病気になった。だから、残る二人で病院中の患者を面倒みなければならなかった。 'Last week, all but two of the nurses at Central Hospital were ill, so those two nurses had to look after all the patients in the hospital.' つまり、二人の看護婦がどの患者も看病した。 Tumari, hutari-no kangohu-ga dono kanzya-mo kanbyou sita. 'In other words, two nurses looked after every patient.'	35
12b	水町診療所は収容患者二十名と小規模だった。四十名の看護婦が働いており、看護婦はそれぞれ、患者一人の面倒をみるので、患者一人につき、二名の看護婦がついていた。 Waterside Nursing Home was a very small care home with only 20 patients. 40 nurses were employed, and each one looked after one patient, so there were two nurses per patient. *(Test sentence as 12a)*	40
12c	*(Context as 12a)* つまり、どの患者も二人の看護婦が看病した。 Tumari, dono kanzya-mo hutari-no kangohu-ga kanbyou sita. 'In other words, two nurses looked after every patient. (*scrambled*)'	17
12d	*(Context as 12b)* *(Test sentence as 12c)*	11
D01	三人の男の人が喫茶店へ出かけた。一人はコーヒーを、一人は紅茶を、そして一人はオレンジジュースを注文した。 'Three boys went to a café. One ordered coffee, one ordered tea and one ordered orange juice.' つまり、だれもが違う飲み物を注文した。 Tumari, daremo-ga tigau nomimono-o tyuumon sita In other words, everyone-NOM different drink-ACC order did 'In other words, everyone ordered something different.'	1
D02	朝日ヶ丘大学の大学生達は先週テストを受けた。最高点は100点満点中90点だった。 'The students at Asahigaoka University sat an exam last week. The highest mark was 90 out of 100.' つまり、だれかが百点満点をとった。 Tumari, dareka-ga hyakutenmanten-o totta In other words, someone-NOM full marks-ACC scored 'In other words, someone scored full marks.'	2

D03	ジェーンとケートは列車でロンドンへ行った。列車に乗っている間中、ジェーンは小説を、ケートは新聞を読んだ。 'Jane and Kate went to London by train. Jane read a novel all the way and Kate read the newspaper.' つまり、新聞をケートさんが読んだ。 Tumari, sinbun-o Keeto-san-ga yonda. In other words, newspaper-ACC Kate-*san*-NOM read 'In other words, the newspaper was read by Kate. (*scrambled*)'	8
D04	ビルが友だちと話していると、ジェーンがぶつかってきた。ビルは倒されてケガをした。 'Bill was chatting to a friend when Jane crashed into him. Bill fell down and hurt himself.' つまり、ビルをジェーンが倒した。 Tumari, Biru-o Jeen-ga taosita. In other words, Bill-ACC Jane-NOM knocked down 'In other words, Jane knocked Bill down. (*scrambled*)'	9
D05	ピーターとジョンは二人ともパーティへワインを持ってきた。ピーターは赤ワイン、ジョンは白ワインだった。 'Peter and John both brought wine to the party. Peter brought red wine and John brought white wine.' つまり、白ワインをピーターが持ってきた。 Tumari, sirowain-o Piitaa-ga mottekita. In other words, white wine-ACC Peter-NOM brought 'In other words, Peter brought the white wine. (*scrambled*)'	38
D06	二人の女の人がハワイへ旅行へ行った。一人は毎日サーフィンをして、そして二人一緒にスキューバダイヴィングをした。 'Two women went on holiday to Hawaii. One of them went surfing every day and both of them went scuba-diving.' つまり、サーフィンを二人とも毎日した。 Tumari, saafin-o hutaritomo mainiti sita. In other words, surfing-ACC both every day did 'In other words, both of them surfed every day. (*scrambled*)'	32
D07	朝日ヶ丘中学校では初年時スペイン語かフランス語のどちらかを選択しなければならない。昨年は全員がスペイン語を選んだ。 'The first year students at Asahigaoka School have to take either Spanish or French. Last year, everyone chose to take Spanish.' つまり、フランス語を数人が選んだ。 Tumari, huransugo-o suunin-ga eranda In other words, French-ACC several people-NOM chose 'In other words, several people chose French. (*scrambled*)'	41

D08	三人の学生が週末ハイキングへ行く約束をした。週末になってみるとそのうちの一人が来られなくなった。 'Three students decided to go hiking at the weekend. When the time came, though, one of them was unable to go.' つまり、だれかが行かなかった。 Tumari, dareka-ga ikanakatta. In other words, someone-NOM did not go 'In other words, someone didn't go.'	4
D09	三人の学生が週末ハイキングへ行く約束をした。週末になってみるとそのうちの一人が来られなくなった。 'Three students decided to go hiking at the weekend. When the time came, though, one of them was unable to go.' つまり、だれもがハイキングに行った。 Tumari, daremo-ga haikingu ni itta. In other words, everyone-NOM hiking to went 'In other words, everyone went hiking.'	37
D10	大学を卒業した後、仲の良かった四人の学生はそれぞれ違う町へ引っ越していった。そのうちの三人は小さなアパートを借り、一人は大きな家を借りた。 'After graduating from university, four friends went to live in different towns. Three of them rented small flats and one of them rented a big house.' つまり、だれかが大きい家を借りた。 Tumari, dareka-ga ookii ie-o karita. In other words, someone-NOM big house-ACC rented 'In other words, someone rented a big house.'	44
D11	大学を卒業した後、仲の良かった四人の学生はそれぞれ違う町へ引っ越していった。そのうちの三人は小さなアパートを借り、一人は大きな家を借りた。 'After graduating from university, four friends went to live in different towns. Three of them rented small flats and one of them rented a big house.' つまり、だれもが大きい家を借りた。 Tumari, daremo-ga ookii ie-o karita. In other words, everyone-NOM big house-ACC rented 'In other words, everyone rented a big house.'	24

D12	朝日ヶ丘大学の事務員の多くはいつも宝くじを買うが、今までだれも当てたことがない。 'Many of the staff in the school office always buy lottery tickets, but so far none of them have ever won.' つまり、だれかがお金をたくさん当てた。 Tumari, dareka-ga okane-o takusan ateta. In other words, someone-NOM money-ACC a lot won 'In other words, someone won lots of money.'	43

Appendix 2E: Tables of results for Phase 2B

Table 2E.1: Mean ratings on test types by participant

id	age	mean rating (scale: –2 to 2)			
		1a	1b	3a	3b
JJ01	20	1	1.75	2	-1.5
JJ02	19	-1.75	-1	2	-2
JJ03	20	-1.75	1.5	2	-0.5
JJ04	22	1.25	1.25	2	1
JJ05	20	1	-0.75	1.75	-1.5
JJ06	21	-0.5	-0.25	1.25	-1
JJ08	22	1	0.33	2	0.5
JJ09	23	0.75	-0.5	0.75	-0.25
JJ10	23	-1	-1.25	1.25	0.5

Table 2E.2: Mean ratings on test tokens

index	mean	SD
1A	-0.6	1.51
1B	0.89	1.45
2A	0	1.76
2B	-0.3	1.70
3A	0.7	1.57
3B	-0.1	1.66
4A	0.2	1.55
4B	0.1	1.52
9A	1.4	0.97
9B	-1	1.22
10A	1.4	0.84
10B	-0.4	1.51
11A	1	0.94
11B	-1.3	0.95
12A	1.5	0.85
12B	-0.9	1.60

2E.3. Repeated measures ANOVA[213]

Table 2E.3.i: Descriptive statistics

	GROUP	Mean	Std. Deviation	N
Type 1a	Phase 2A (picture context)	1.5100	.35707	25
	Phase 2B (written context)	.0750	1.20214	10
Type 1b	Phase 2A (picture context)	-.9667	.74068	25
	Phase 2B (written context)	.1333	1.07022	10
Type 3a	Phase 2A (picture context)	1.5700	.47610	25
	Phase 2B (written context)	1.3250	.33437	10
Type 3b	Phase 2A (picture context)	-1.0900	.95982	25
	Phase 2B (written context)	-.9167	.84163	10

Table 2E.3.ii: Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Subj QP	Scope	Type III Sum of Squares	df	Mean Square	F	Sig.
Subj QP	Linear		3.335E-02	1	3.335E-02	.063	.803
Subj QP * Group	Linear		.124	1	.124	.235	.631
Error(Subj QP)	Linear		17.377	33	.527		
Scope		Linear	95.683	1	95.683	137.063	.000
Scope * Group		Linear	15.575	1	15.575	22.311	.000
Error(Scope)		Linear	23.037	33	.698		
Subj QP * Scope	Linear	Linear	11.012	1	11.012	45.051	.000
Subj QP * Scope * Group	Linear	Linear	8.000	1	8.000	32.729	.000
Error(Subj QP*Scope)	Linear	Linear	8.067	33	.244		

Table 2E.3.iii: Tests of Between-Subjects Contrasts

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	4.803	1	4.803	6.098	.019
GROUP	.295	1	.295	.375	.545
Error	25.992	33	.788		

[213] Sphericity is assumed, as Mauchly's test of sphericity was non-significant. Equality of variance is not assumed, as Levene's test was significant for Type 1a (see Table 2E.4.i).

2E.4. Individual samples *t*-tests

Table 2E.4.i. Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
Type 1A	51.290	1	33	.000
Type 1B	3.018	1	33	.092
Type 3A	4.065	1	33	.052
Type 3B	.762	1	33	.389

Table 2E.4.ii. T-test comparisons of Phase 2A and Phase 2B results for Types 1a, 1b, 3a, and 3b

Note: Based on the results of Levene's Test, (Table 2E.4.i) the 'Equal variances assumed' figures are used for Types 1b, 3a and 3b; and the 'Equal variances not assumed' figures for Type 1a.

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Type 1a	Equal variances assumed	5.497	33	.000	1.43500	.26107	.90384	1.96616
	Equal variances not assumed	3.710	9.642	.004	1.43500	.38680	.56880	2.30120
Type 1b	Equal variances assumed	-3.486	33	.001	-1.10000	.31558	-1.7421	-.45795
	Equal variances not assumed	-2.978	12.606	.011	-1.10000	.36943	-1.9006	-.29934
Type 3a	Equal variances assumed	1.482	33	.148	.24500	.16537	-.09145	.58145
	Equal variances not assumed	1.722	23.677	.098	.24500	.14229	-.04889	.53889
Type 3b	Equal variances assumed	-.499	33	.621	-.17333	.34763	-.88059	.53392
	Equal variances not assumed	-.528	18.884	.604	-.17333	.32815	-.86045	.51378

Appendix 3: Participant data

Table 3.i: Data on the main study (Phase 3) participants
Note: Missing data is coded as '99'.

	id	age	gender	proficiency task score (out of 42)	dialect (native Chinese speakers)
native Japanese	JJ01	99	m	24	n/a
	JJ02	28	f	25	n/a
	JJ03	27	m	28	n/a
	JJ04	25	f	26	n/a
	JJ05	23	f	33	n/a
	JJ06	24	f	24	n/a
	JJ07	24	f	24	n/a
	JJ08	23	f	26	n/a
	JJ09	24	m	32	n/a
	JJ10	31	m	24	n/a
	JJ11	20	f	20	n/a
	JJ12	20	f	20	n/a
	JJ13	19	f	19	n/a
	JJ14	20	f	17	n/a
	JJ15	20	f	25	n/a
	JJ16	19	f	23	n/a
	JJ17	20	f	23	n/a
	JJ18	19	f	12	n/a
	JJ19	36	m	n/a	n/a
	JJ20	57	m	n/a	n/a
	JJ22	21	m	n/a	n/a
	JJ23	21	f	n/a	n/a
	JJ24	21	f	n/a	n/a
	JJ25	20	m	n/a	n/a
	JJ26	20	m	n/a	n/a
	JJ27	23	f	n/a	n/a
	JJ28	22	m	n/a	n/a
	JJ29	21	m	n/a	n/a
	JJ30	21	f	n/a	n/a
	JJ31	20	m	n/a	n/a
	JJ32	20	f	n/a	n/a
	JJ33	20	f	n/a	n/a
	JJ34	20	f	n/a	n/a
	JJ35	20	f	n/a	n/a
	JJ36	23	m	n/a	n/a
	JJ37	20	m	n/a	n/a
	JJ38	21	m	n/a	n/a
	JJ39	20	m	n/a	n/a
	JJ40	21	m	n/a	n/a
	JJ101	17	f	19	n/a
	JJ102	18	f	18	n/a
	JJ103	19	f	23	n/a
	JJ104	18	f	19	n/a
	JJ105	18	f	25	n/a
	JJ106	17	f	20	n/a
	JJ107	18	f	22	n/a
	JJ108	18	f	23	n/a
	JJ109	18	f	20	n/a

	id	age	gender	proficiency task score (out of 42)	dialect (native Chinese speakers)
	JJ110	18	f	20	*n/a*
	JJ111	18	f	15	*n/a*
	JJ112	18	f	23	*n/a*
L1 Chinese/L2 Japanese	CJ01	29	f	11	Mandarin
	CJ02	27	f	17	Mandarin
	CJ03	33	f	12	Mandarin
	CJ04	34	f	11	Mandarin + Shanghai Dialect
	CJ05	27	f	20	Mandarin
	CJ06	29	f	17	Mandarin
	CJ07	25	f	15	Taiwanese
	CJ08	29	f	14	Taiwanese
	CJ09	24	f	13	Mandarin
	CJ10	25	f	3	Mandarin
	CJ11	23	f	4	Mandarin
	CJ12	21	f	3	Mandarin
	CJ13	24	f	10	Mandarin
	CJ14	20	f	6	Mandarin
	CJ15	25	m	12	Cantonese
	CJ16	21	f	14	Cantonese
	CJ18	24	f	13	Cantonese
L1 English/L2 Japanese	EJ02	20	f	8	*n/a*
	EJ03	20	f	6	*n/a*
	EJ04	21	f	15	*n/a*
	EJ05	21	m	12	*n/a*
	EJ06	21	m	14	*n/a*
	EJ07	22	m	11	*n/a*
	EJ08	21	m	8	*n/a*
	EJ09	22	f	8	*n/a*
	EJ10	21	m	18	*n/a*
	EJ11	20	m	10	*n/a*
	EJ12	22	m	14	*n/a*
	EJ13	19	m	4	*n/a*
	EJ14	21	m	8	*n/a*
	EJ15	23	m	16	*n/a*
	EJ16	21	f	7	*n/a*
	EJ17	19	f	5	*n/a*
	EJ18	19	m	3	*n/a*
	EJ19	24	m	7	*n/a*
	EJ20	21	f	15	*n/a*
	EJ21	22	m	8	*n/a*
	EJ22	21	f	9	*n/a*
	EJ23	23	m	12	*n/a*
	EJ24	22	m	12	*n/a*
	EJ25	21	m	7	*n/a*
	EJ26	21	m	8	*n/a*
	EJ27	20	m	8	*n/a*
	EJ28	20	m	4	*n/a*
	EJ29	30	m	8	*n/a*
	EJ30	22	f	4	*n/a*
L1 Korean/L2 Japanese	KJ01	21	f	20	*n/a*
	KJ02	21	f	16	*n/a*
	KJ03	27	f	25	*n/a*
	KJ04	27	f	7	*n/a*
	KJ05	26	f	8	*n/a*
	KJ06	21	f	13	*n/a*

	id	age	gender	proficiency task score (out of 42)	dialect (native Chinese speakers)
L1 Korean/L2 Japanese *(cont.)*	KJ07	21	f	29	*n/a*
	KJ08	31	f	17	*n/a*
	KJ09	21	f	17	*n/a*
	KJ10	21	f	22	*n/a*
	KJ11	27	f	10	*n/a*
	KJ12	28	f	16	*n/a*
	KJ13	27	f	14	*n/a*
	KJ14	25	m	2	*n/a*
	KJ15	25	m	9	*n/a*
	KJ16	26	m	3	*n/a*
	KJ17	24	f	11	*n/a*
	KJ18	26	m	7	*n/a*
	KJ19	26	m	1	*n/a*
	KJ20	22	f	7	*n/a*
	KJ21	22	f	7	*n/a*
	KJ22	29	f	17	*n/a*
	KJ23	24	f	12	*n/a*
	KJ25	27	m	6	*n/a*
	KJ26	23	f	7	*n/a*
	KJ27	22	f	9	*n/a*
	KJ29	25	m	2	*n/a*
	KJ30	23	f	9	*n/a*
	KJ31	24	f	9	*n/a*
	KJ32	35	f	8	*n/a*
	KJ33	33	f	8	*n/a*
	KJ34	22	f	23	*n/a*
	KJ35	22	f	6	*n/a*
	KJ36	23	f	12	*n/a*
	KJ38	27	f	17	*n/a*
	KJ39	51	f	4	*n/a*
	KJ41	69	f	9	*n/a*
	KJ42	24	f	7	*n/a*
native English	EE01	18	f	*n/a*	*n/a*
	EE02	23	f	*n/a*	*n/a*
	EE03	20	f	*n/a*	*n/a*
	EE04	18	m	*n/a*	*n/a*
	EE05	18	f	*n/a*	*n/a*
	EE06	18	m	*n/a*	*n/a*
	EE07	18	f	*n/a*	*n/a*
	EE08	18	f	*n/a*	*n/a*
	EE09	19	f	*n/a*	*n/a*
	EE10	19	f	*n/a*	*n/a*
	EE11	18	f	*n/a*	*n/a*
	EE12	18	f	*n/a*	*n/a*
	EE13	18	f	*n/a*	*n/a*
	EE14	18	f	*n/a*	*n/a*
	EE15	18	f	*n/a*	*n/a*
	EE16	18	f	*n/a*	*n/a*
	EE17	18	f	*n/a*	*n/a*
	EE18	19	m	*n/a*	*n/a*
	EE19	18	f	*n/a*	*n/a*
	EE20	18	f	*n/a*	*n/a*
	EE21	19	f	*n/a*	*n/a*
	EE22	18	f	*n/a*	*n/a*
	EE23	18	f	*n/a*	*n/a*

	id	age	gender	proficiency task score (out of 42)	dialect (native Chinese speakers)
	EE24	18	f	*n/a*	*n/a*
	EE16	18	f	*n/a*	*n/a*
	EE25	19	f	*n/a*	*n/a*
	EE26	18	f	*n/a*	*n/a*
	EE27	18	m	*n/a*	*n/a*
	EE28	18	f	*n/a*	*n/a*
	EE29	19	f	*n/a*	*n/a*
	EE30	19	f	*n/a*	*n/a*
	EE31	18	f	*n/a*	*n/a*
	EE32	18	f	*n/a*	*n/a*
	EE33	24	f	*n/a*	*n/a*
	EE34	20	m	*n/a*	*n/a*
	EE36	19	f	*n/a*	*n/a*
	EE37	18	f	*n/a*	*n/a*
	EE38	18	f	*n/a*	*n/a*
	EE39	19	f	*n/a*	*n/a*
	EE40	20	f	*n/a*	*n/a*
native Chinese	CC01	33	f	*n/a*	Mandarin
	CC02	32	m	*n/a*	Mandarin
	CC03	34	f	*n/a*	Mandarin
	CC04	30	f	*n/a*	Mandarin
	CC05	32	m	*n/a*	Mandarin
	CC06	27	m	*n/a*	Mandarin
	CC07	31	m	*n/a*	Mandarin
	CC08	30	f	*n/a*	Mandarin
	CC09	33	m	*n/a*	Mandarin
	CC10	30	m	*n/a*	Mandarin
	CC11	32	m	*n/a*	Mandarin
	CC12	30	f	*n/a*	Shanghai Dialect
	CC13	33	m	*n/a*	Mandarin
	CC14	35	m	*n/a*	Mandarin & Cantonese
	CC15	36	f	*n/a*	Mandarin
	CC16	31	f	*n/a*	Mandarin
	CC17	21	m	*n/a*	Cantonese
	CC18	23	m	*n/a*	Mandarin
	CC19	29	f	*n/a*	Mandarin & Taiwanese
	CC20	25	m	*n/a*	Mandarin
	CC21	22	f	*n/a*	Mandarin
	CC22	29	f	*n/a*	Mandarin & Taiwanese
	CC23	24	m	*n/a*	Mandarin
	CC24	22	m	*n/a*	Mandarin
	CC25	26	f	*n/a*	Mandarin
	CC26	22	m	*n/a*	Mandarin
	CC27	27	f	*n/a*	Mandarin
	CC28	24	f	*n/a*	Taiwanese
	CC29	25	f	*n/a*	Mandarin
	CC30	21	m	*n/a*	Mandarin
	CC31	20	f	*n/a*	Mandarin
	CC32	20	f	*n/a*	Mandarin
	CC33	21	f	*n/a*	Mandarin
	CC34	35	f	*n/a*	Mandarin
	CC35	24	f	*n/a*	Mandarin
	CC36	23	f	*n/a*	Mandarin
	CC37	22	m	*n/a*	Shanghai Dialect & Mandarin

	id	age	gender	proficiency task score (out of 42)	dialect (native Chinese speakers)
native Korean	KK01	21	f	*n/a*	*n/a*
	KK02	22	f	*n/a*	*n/a*
	KK03	21	f	*n/a*	*n/a*
	KK04	21	f	*n/a*	*n/a*
	KK05	23	f	*n/a*	*n/a*
	KK07	21	f	*n/a*	*n/a*
	KK08	21	f	*n/a*	*n/a*
	KK09	21	f	*n/a*	*n/a*
	KK10	25	m	*n/a*	*n/a*
	KK11	24	m	*n/a*	*n/a*
	KK12	21	f	*n/a*	*n/a*
	KK13	25	m	*n/a*	*n/a*
	KK14	20	f	*n/a*	*n/a*
	KK15	21	f	*n/a*	*n/a*
	KK16	26	m	*n/a*	*n/a*
	KK17	20	f	*n/a*	*n/a*
	KK18	21	f	*n/a*	*n/a*
	KK19	24	m	*n/a*	*n/a*
	KK20	21	m	*n/a*	*n/a*
	KK21	25	m	*n/a*	*n/a*
	KK22	27	m	*n/a*	*n/a*
	KK23	*99*	f	*n/a*	*n/a*
	KK24	21	f	*n/a*	*n/a*
	KK25	21	f	*n/a*	*n/a*
	KK26	22	f	*n/a*	*n/a*

Appendix 4: Proficiency task

About this appendix

This appendix comprises the following two sections:

1. Appendix 4A: the cloze test used as the proficiency task in Phase 3
2. Appendix 4B: statistical analyses of proficiency task scores

Appendix 4A: Phase 3 proficiency task

The actual test is on the following two pages. The English translation below is provided for reference (adapted from *Nihongo Journal* 2000.5: 46–47). It was not used in the actual proficiency test.

4A.i. English translation of the proficiency test passage:

Room-sharing

'Room-sharing' means renting a condominium or apartment communally with friends and acquaintances instead of living with family members or siblings. This is a common practice in Europe and North America and now it is also starting to become popular among young people in Japanese cities. They like room-sharing for a variety of reasons: it lets you save money on rent, it's fun, and it provides a sense of security. More and more people are learning about these advantages by experiencing room-sharing while studying abroad, and this has lowered resistance to the idea.

Living with friends

Ms F (27) once shared a room with a woman from Hong Kong while living in China. Now, she shares a two-bedroom apartment in Tokyo with two friends from her university days. Each of the three friends pays 40,000 yen to cover the rent of 120,000 yen. They each also put 30,000 yen into a common purse every month to buy food and pay the electricity, water, telephone and other utility bills. After beginning this shared living arrangement, Ms F said that the cost-of-living savings were the biggest advantage. The rent on an apartment for a single person in Tokyo is 60,000 to 80,000 yen, and now she's only paying half of that. The utilities are also much lower than she would have to pay living alone. One disadvantage she mentioned was that she usually can't talk for a long time on the telephone.

Finding roommates on the Internet

An increasing number of people are finding roommates on the Internet. Many Japanese people want to find a foreign roommate so they can learn a foreign language while sharing accommodation. Last August, borderless-tokyo-corp (an international exchange association) established a website where people looking for roommates can exchange information (http://borderless-tokyo.com). The bulletin board for people seeking roommates is full of ads from both Japanese and foreigners. According to Seiji Kondo, who manages the website, the site sometimes gets as many as 200 hits a day. However, Kondo mentioned that landlords are resistant to the idea of room-sharing, and places that allow it can be few and far between.

4A.ii. Proficiency test:

以下のテキストの空白（……………）に適切な言葉を書いて下さい。

例：

- 雪がたくさん降っている(i)……………、飛行機は飛ばないかもしれない。
- 新しい携帯(ii)……………は軽くて便利です。
- 窓を閉めろと言ったのに、田中君(iii)……………開けたまま教室を出て(iv)……………。

* *

「ルームシェア」

マンションやアパートの一部屋を家族や兄弟とではなく、友人同士などで借りて共同生活をすることを「ルームシェア」という。欧米では一般的だ(1)……………、日本でも最近、都市部の(2)……………の間で広まっている。「家賃が(3)……………できる」「楽しい」「安心」など、理由は(4)……………だが、海外留学などで経験(5)……………ルームシェアの利点を知る人(6)……………増え、他人と生活すること(7)……………の抵抗感がなくなってきて(8)……………、という背景もあるよう(9)……………

友だちと一緒に暮らす

Ｆさん（27歳・女性）は中国で、香港(10)……………の女性とルームシェアを経験(11)……………。現在も都内で大学時代(12)……………同級生の女性２人と2LDK(13)……………部屋をルームシェアしている。家賃12万(14)……………は３人で４万円ずつ(15)……………しているという。電気、水道、

1

(16)……………、などの公共料金と食費(17)……………、3人共通の財布を用意(18)……………毎月１人３万円ずつ入れ、(19)……………から支払う。

住みはじめてからの(20)……………を聞いてみると、一番の(21)……………はやはり金銭面。都内で１人(22)……………住むには６万～８万円かかる(23)……………、今はその半分。公共料金(24)……………は１人で負担するより(25)……………安くてすむ。一方、デメリットは、長電話(26)……………しにくいことだという。

インターネットで「シェアメート」を探す

(27)……………をインターネットを通して探す人(28)……………増えている。また、「一緒に暮らして(29)……………を覚えたい」という理由から、(30)……………に外国人を希望する日本人(31)……………多い。

「国際交流協会」は、シェアメート(32)……………探す人たちの情報交換の(33)……………を設けようと、昨年８月に(34)……………を立ち上げた（http://borderless-tokyo.com）。シェアメート募集の(35)……………には、日本人、外国人から多数(36)……………書き込みがある。ホームページの管理者、近藤誠二(37)……………によると、最近はアクセス(38)……………が多いときは１日に200(39)……………になるという。だが、(40)……………一般には、大家さんの(41)……………が得られず、シェアを受け入れてくれる(42)……………は少ないそうだ。

2

A.iii. Cloze test answers

Possible answers to the example items:

から／ので

i. 電話

ii. は

v. いった

Answers to cloze test items:

1. が
2. 若者
3. 節約
4. さまざま
5. して
6. が
7. へ
8. いる
9. だ
10. 出身
11. した
12. の
13. の
14. 円
15. 負担
16. 電話
17. は
18. して
19. そこ
20. 感想
21. メリット
22. で
23. が
24. など
25. ずっと
26. が
27. シェアメート
28. も
29. 外国語
30. 相手
31. も
32. を
33. 場
34. ホームページ
35. 掲示板
36. の
37. さん
38. 数
39. 近く
40. まだ
41. 理解
42. ところ

Appendix 4B: Statistical analyses of proficiency task scores

Table 4B.1: Descriptive Statistics

GROUP	Mean score	Std. Deviation	N
EJ int	7.05	2.114	20
EJ adv	14.22	2.048	9
CJ int	6.86	3.716	7
CJ adv	14.70	2.584	10
KJ int	6.78	2.696	23
KJ adv	18.00	4.928	15

Table 4B.2: Individual samples *t*-tests

Note: Based on the results of Levene's Test, the 'Equal variances assumed' figures are used in Chapter 5 for the comparisons between the EJ intermediate and advanced groups and between the CJ intermediate and advanced groups; the 'Equal variances not assumed' figures are used for the comparison between the KJ intermediate and advanced groups.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
EJ int v. adv	Equal variances assumed	.017	.898	-8.529	27	.000	-7.17	.841	-8.898	-5.447
	Equal variances not assumed			-8.637	15.969	.000	-7.17	.830	-8.933	-5.412
CJ int v. adv	Equal variances assumed	3.472	.082	-5.155	15	.000	-7.84	1.521	-11.086	-4.600
	Equal variances not assumed			-4.826	9.987	.001	-7.84	1.625	-11.464	-4.222
KJ int v. adv	Equal variances assumed	6.056	.019	-9.070	36	.000	-11.22	1.237	-13.726	-8.709
	Equal variances not assumed			-8.064	19.526	.000	-11.22	1.391	-14.124	-8.311

Appendix 5: QP-QP test

About this appendix

This appendix relates to the investigation of doubly-quantified sentences in the main experimental study (Phase 3). It should be used in conjunction with Chapter 5, Section 5.2. The appendix comprises the following sections:

1. Appendix 5A.i: Test pictures and sentences in Japanese
   Appendix 5A.ii: Distractor pictures and sentences in Japanese
2. Appendix 5B: Test sentences and distractors in Chinese and English
3. Appendix 5C: Test sentences and distractors in Korean
4. Appendix 5D: L2ers' answer sheet in English
5. Appendix 5E: L2ers' answer sheet in Japanese
6. Appendix 5F: Tables of results

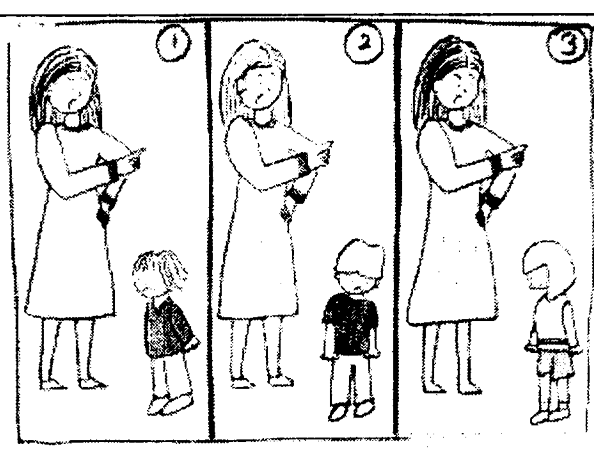
Appendix 5A.i: Test pictures and sentences in Japanese

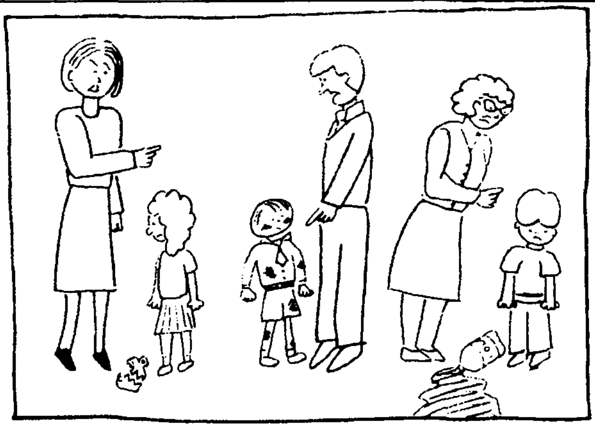
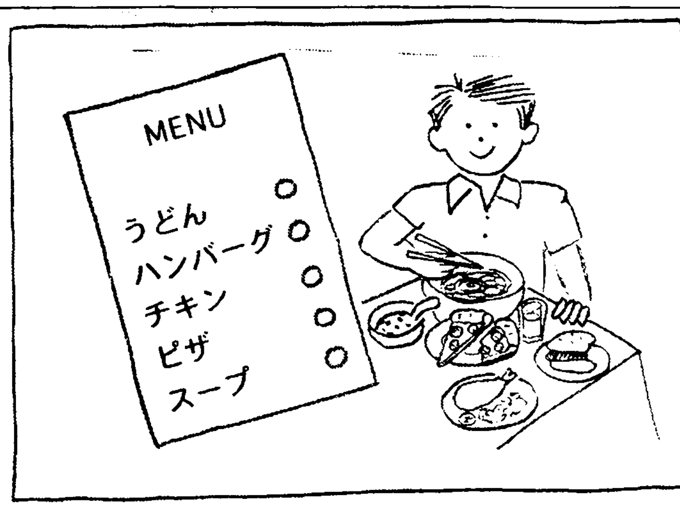
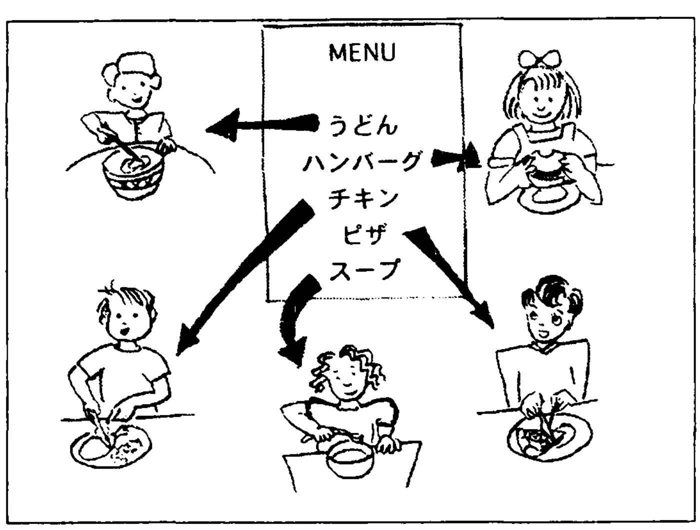
Notes:

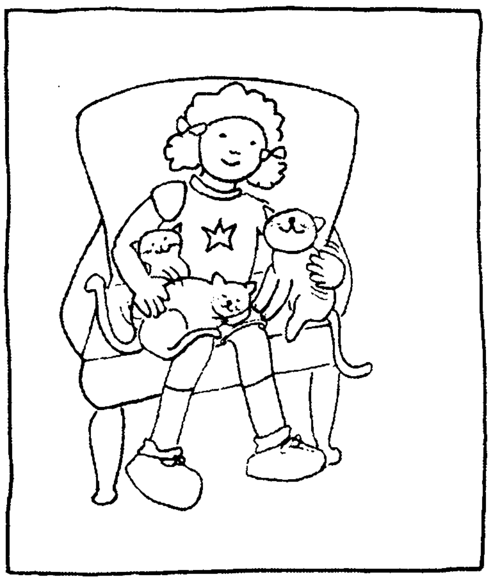
i. The test type can be identified by the item index. For example, index '1.1a' is Type 1a, Token 1; '1.2a' is Type 1a, Token 2, etc.

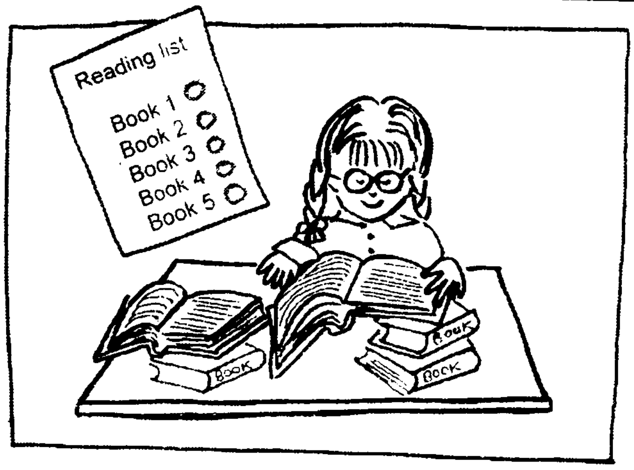
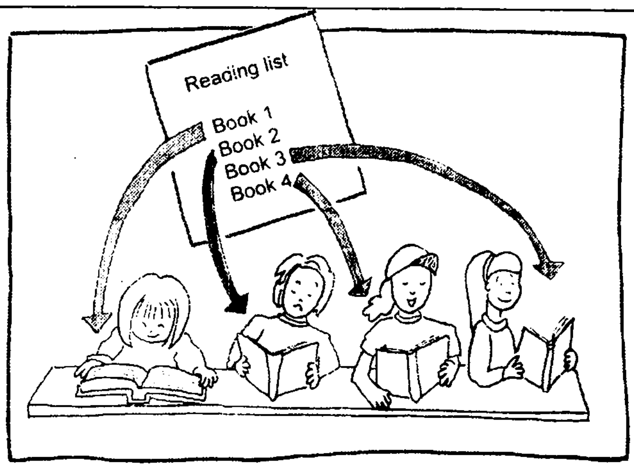
ii. The 'Item no.' columns give the number of each item in the two test orders. 'T1' indicates 'QP-QP Task 1'; 'T2' indicates 'QP-QP Task 2'. Thus, for example, index 1.1a was item no. 2 in Order 1 of QP-QP Task 1, and item no. 43 in Order 2 of QP-QP Task 1.

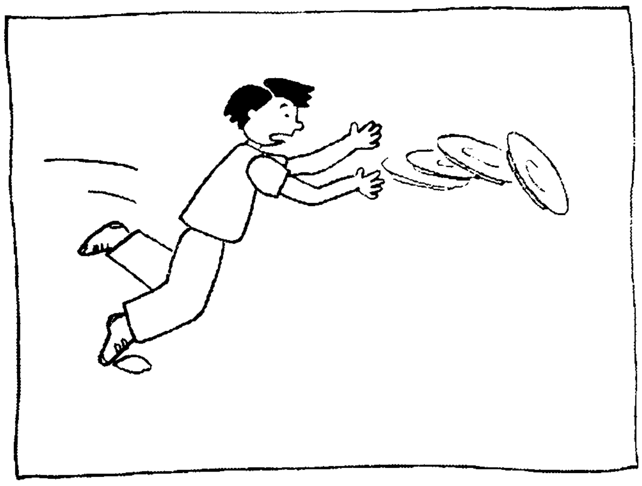
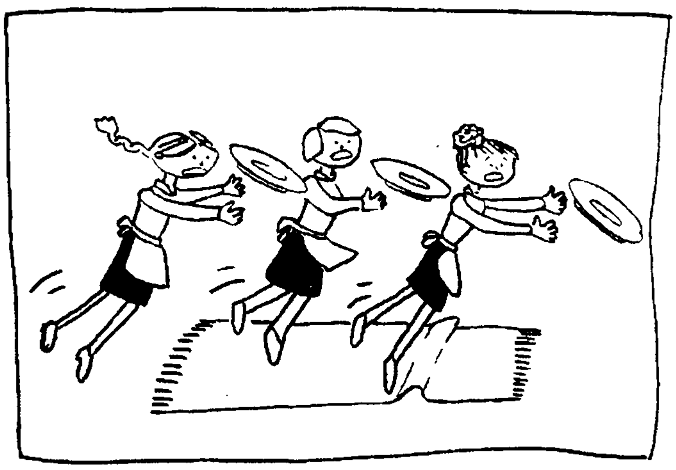
iii. Colour was used to facilitate comprehension of the pictures. Please contact the author about viewing the pictures in colour, if necessary.

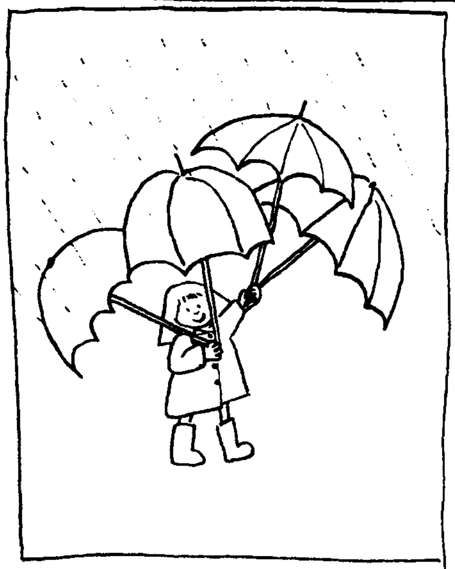
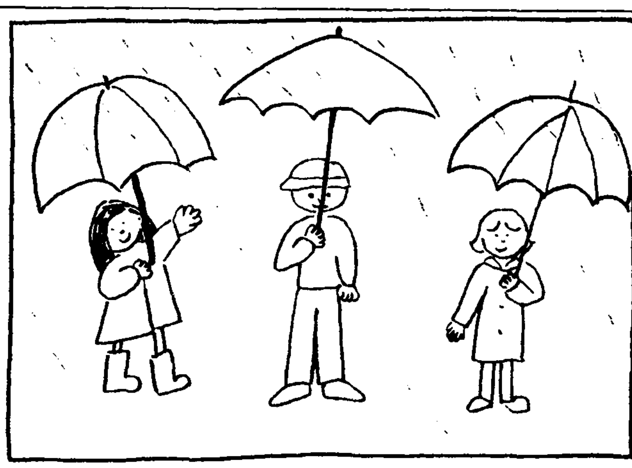
index	picture & sentence	item no.	
		order 1	order 2
1.1a	 だれかがどの子供も叱った。 Dareka-ga dono kodomo-mo sikatta. someone-NOM every child-QPt scolded 'Someone scolded every child.'	T1: 2	T1: 43

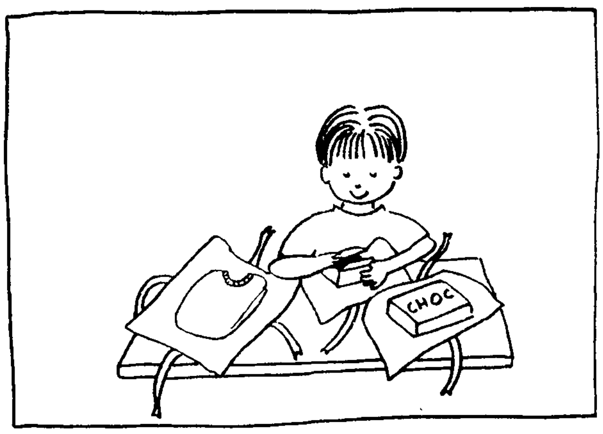
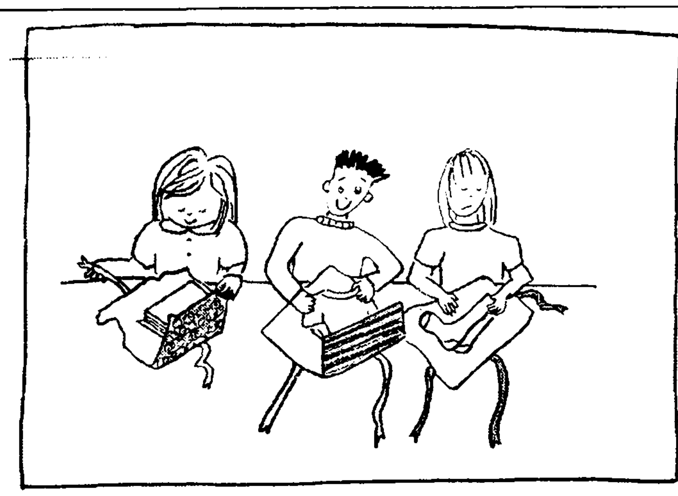
1.1b	 (*Sentence as 1.1a*)	T1: 29	T1: 16
1.1c	*(Picture as 1.1a)* どの子供もだれかが叱った。 Dono kodomo-mo dareka-ga sikatta. every child-QPt someone-NOM scolded 'Someone scolded every child. (*scrambled*)'	T2: 39	T2: 6
1.1d	*(Picture as 1.1b)* (*Sentence as 1.1c*)	T2: 18	T2: 27
1.2a	 だれかがどの料理も食べてみた。 Dareka-ga dono ryouri-mo tabete-mita. someone-NOM every dish-QPt tried 'Someone tried every dish.'	T1: 8	T1: 37
1.2b	 (Sentence as 1.2a)	T1: 12	T1: 33

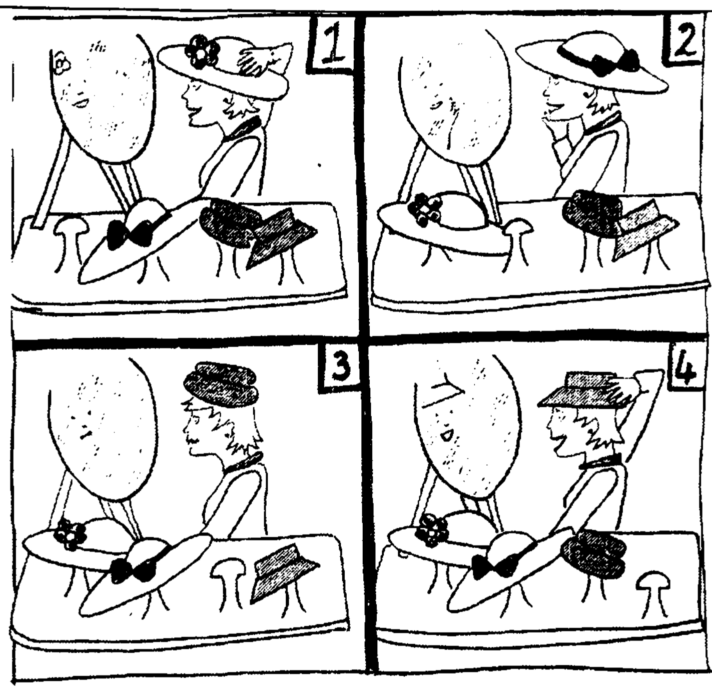
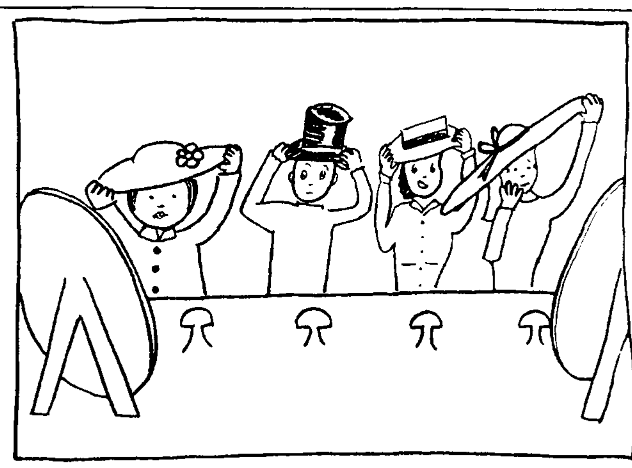
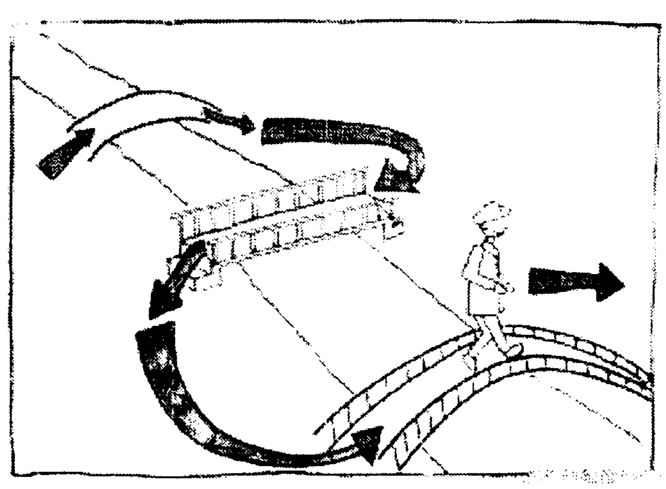
1.2c	*(Picture as 1.2a)* どの料理もだれかが食べてみた。 Dono ryouri-mo dareka-ga tabete-mita. every dish-QPt someone-NOM tried 'Someone tried every dish. (*scrambled*)'	T2: 5	T2: 40
1.2d	*(Picture as 1.2b)* (*Sentence as 1.2c*)	T2: 17	T2: 28
1.3a	 だれかがどの猫もなでた。 Dareka-ga dono neko-mo nadeta. someone-NOM every cat-QPt stroked 'Someone stroked every cat.'	T1: 26	T1: 19
1.3b	 (Sentence as 1.3a)	T1: 23	T1: 22
1.3c	*(Picture as 1.3a)* どの猫もだれかがなでた。 Dono neko-mo dareka-ga nadeta. every cat-QPt someone-NOM stroked 'Someone stroked every cat. (*scrambled*)'	T2: 26	T2: 19
1.3d	*(Picture as 1.3b)* (*Sentence as 1.3c*)	T2: 40	T2: 5

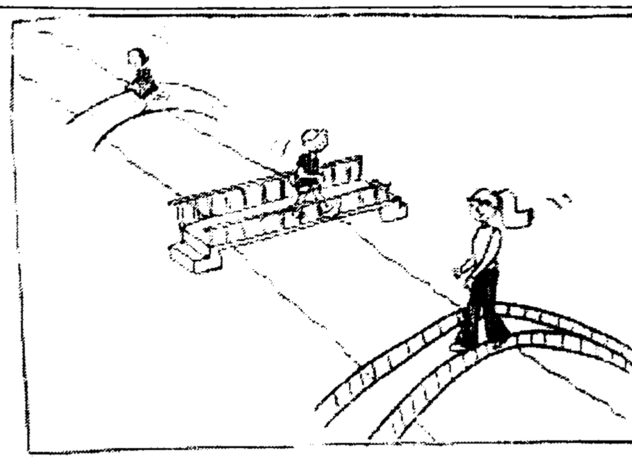
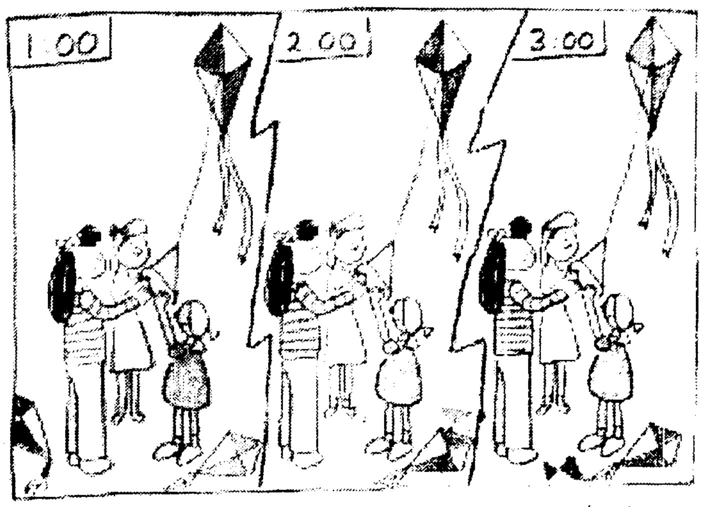
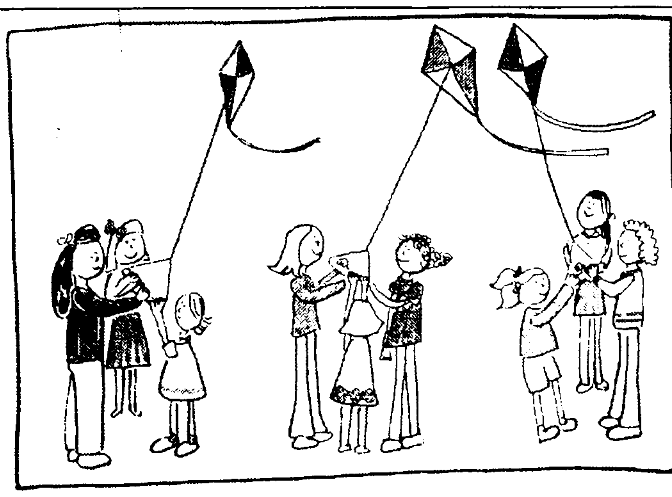
1.4a	 だれかがどの本も読んだ。 Dareka-ga dono hon-mo yonda. someone-NOM every book-QPt read 'Someone read every book.	T1: 13	T1: 32
1.4b	 (*Sentence as 1.4a*)	T1: 42	T1: 3
1.4c	*(Picture as 1.4a)* どの本もだれかが読んだ。 Dono hon-mo dareka-ga yonda. every book-QPt someone-NOM read 'Someone read every book. (*scrambled*)'	T2: 3	T2: 42
1.4d	*(Picture as 1.4b)* (*Sentence as 1.4c*)	T2: 34	T2: 11

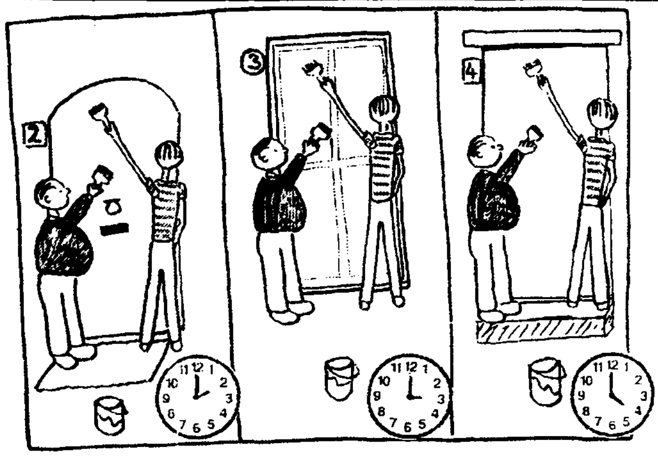
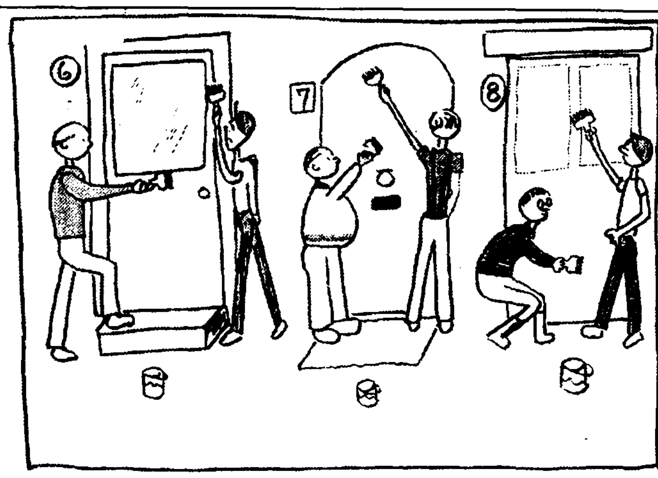
1.5a	 だれかがどの皿も落とした。 Dareka-ga dono sara-mo otosita. someone-NOM every plate-QPt dropped 'Someone dropped every plate.'	T1: 35	T1: 10
1.5b	 (*Sentence as 1.5a*)	T1: 5	T1: 40
1.5c	*(Picture as 1.5a)* どの皿もだれかが落とした。 Dono sara-mo dareka-ga otosita. every plate-QPt someone-NOM dropped 'Someone dropped every plate. (*scrambled*)'	T2: 19	T2: 26
1.5d	*(Picture as 1.5b)* (*Sentence as 1.5c*)	T2: 32	T2: 13

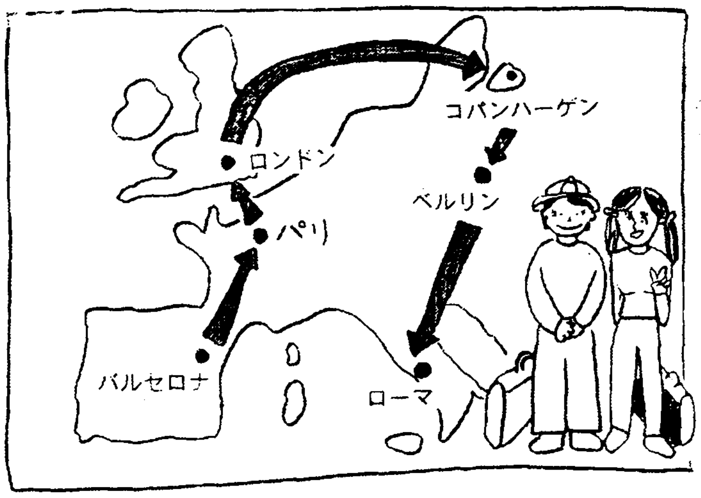
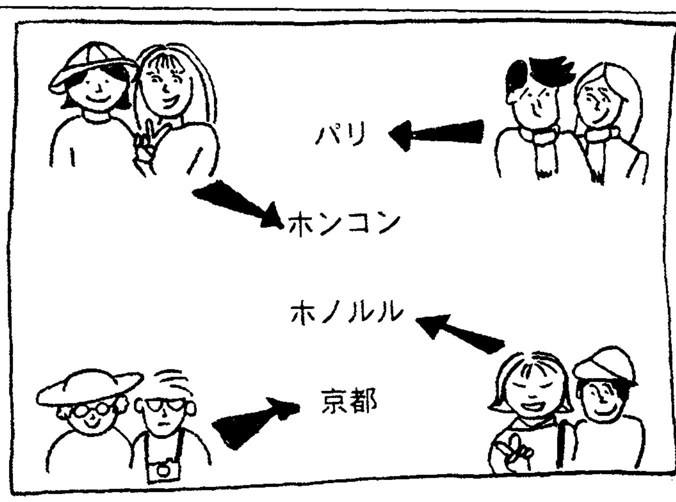
2.1a	 だれかがすべての傘をさした。 Dareka-ga subete-no kasa-o sasita. someone-NOM all-GEN umbrella-ACC put up 'Someone put up all the umbrellas.'	T2: 6	T2: 39
2.1b	 (*Sentence as 2.1a*)	T2: 30	T2: 15
2.2a	 だれかがすべてのスーツケースを運んだ。 Dareka-ga subete-no suutukeesu-o hakonda. someone-NOM all-GEN suitcase-ACC carried 'Someone carried all the suitcases.'	T2: 35	T2: 10

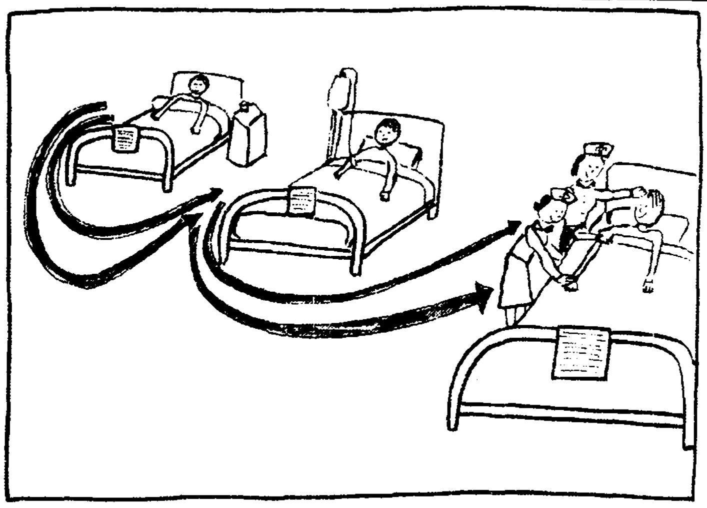
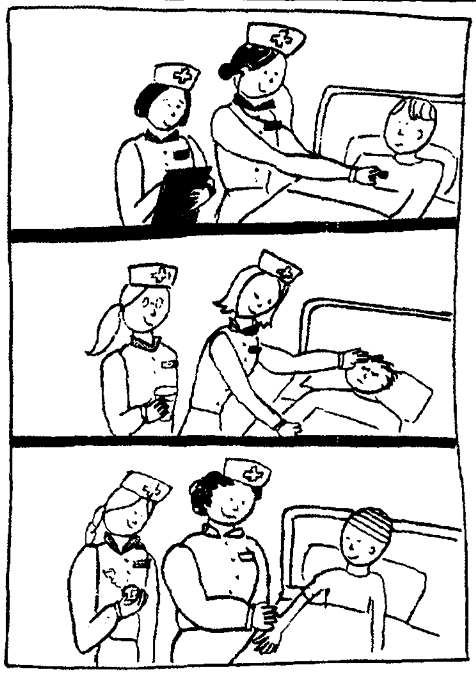
2.2b	 (Sentence as 2.2a)	T2: 27	T2: 18
2.3a	 だれかがすべてのプレゼントを開けた。 Dareka-ga subete-no purezento-o aketa. someone-NOM all-GEN present-ACC opened 'Someone opened all the presents.'	T2: 7	T2: 38
2.3b	 (Sentence as 2.3a)	T2: 11	T2: 34

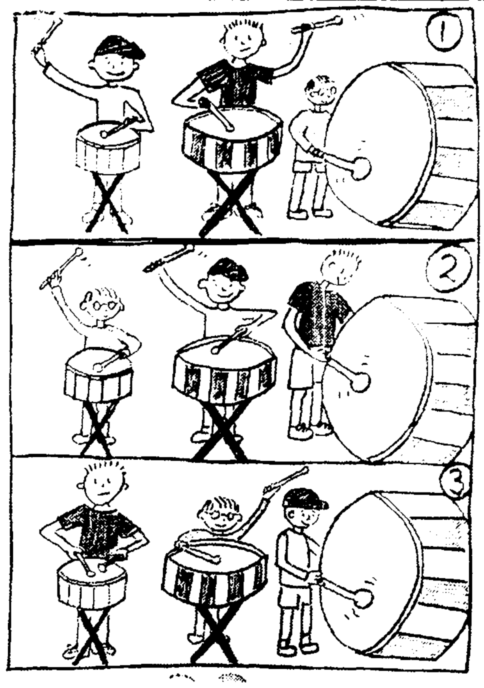
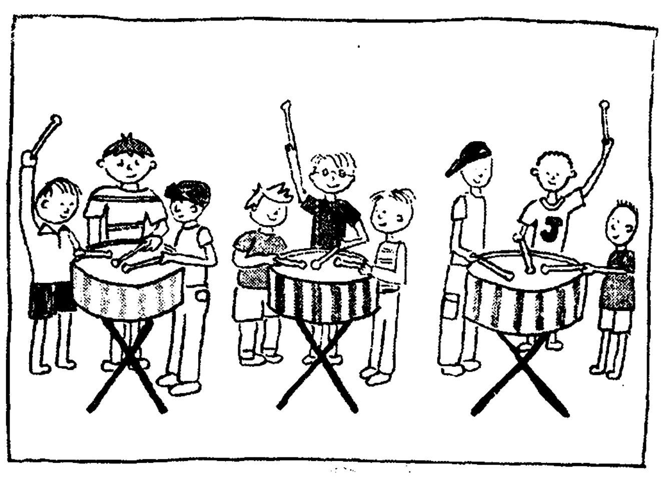
2.4a	 だれかがすべての帽子を被ってみた。 Dareka-ga subete-no bousi-o kabutte-mita. someone-NOM all-GEN bousi-ACC tried on 'Someone tried on all the hats.'	T2: 36	T2: 9
2.4b	 (Sentence as 8a)	T2: 13	T2: 32
2.5a	 だれかがすべての橋を渡った。 Dareka-ga subete-no hasi-o watatta. someone-NOM all-GEN hasi-ACC crossed 'Someone crossed all the bridges.'	T2: 10	T2: 35

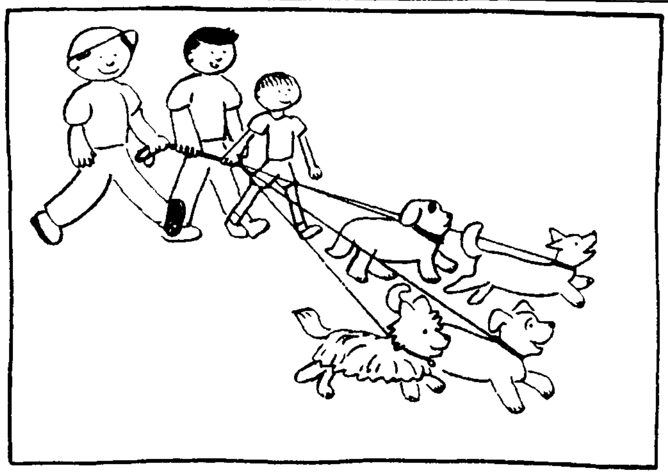
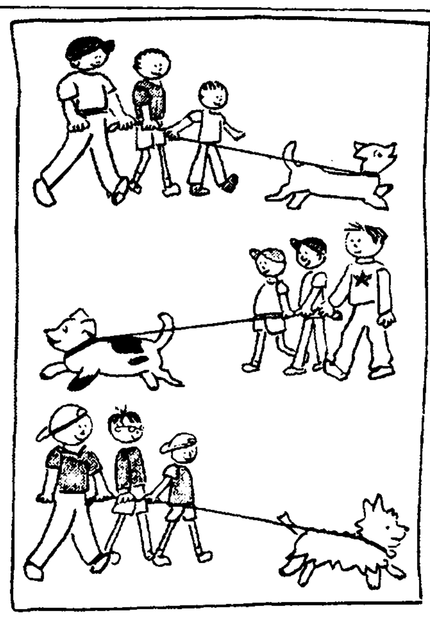
2.5b	 (Sentence as 2.5a)	T2: 25	T2: 20
3.1a	 三人の女の子がどのたこも上げた。 Sannin-no onnanoko-ga dono tako-mo ageta. three.HUMAN-GEN girl-NOM every kite-QPt flew 'Three girls flew every kite.'	T2: 2	T2: 43
3.1b	 (*Sentence as 3.1a*)	T2: 43	T2: 2
3.1c	*(Picture as 3.1a)* どのたこも三人の女の子が上げた。 Dono tako-mo sannin-no onnanoko-ga ageta. every kite- QPt three.HUMAN-GEN girl-NOM flew 'Three girls flew every kite.' (*scrambled*)'	T1: 16	T1: 29

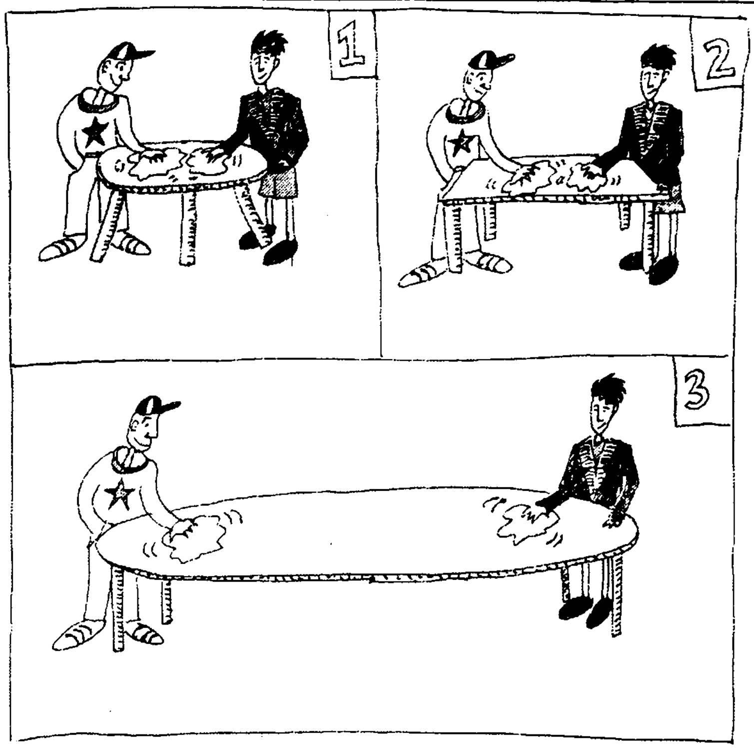
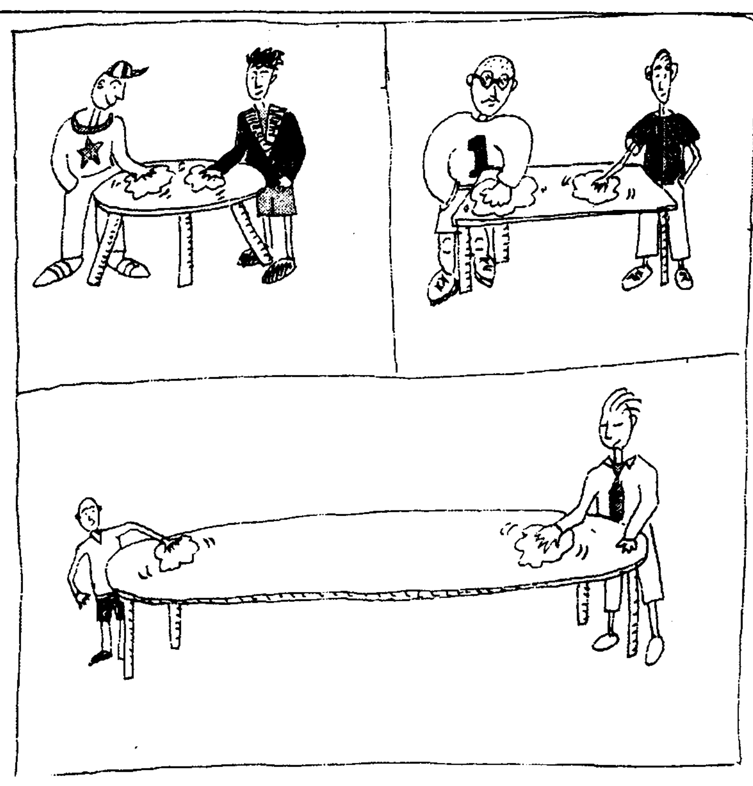
3.1d	*(Picture as 3.1b)* (*Sentence as 3.1c*)	T1: 22	T1: 23
3.2a	 二人の男がどのドアもぬった。 Hutari-no otoko-ga dono doa-mo nutta. two.HUMAN-GEN man-NOM every door-QPt painted 'Two men painted every door.'	T2: 42	T2: 3
3.2b	 (*Sentence as 3.2a*)	T2: 28	T2: 17
3.2c	*(Picture as 3.2a)* どのドアも二人の男がぬった。 Dono doa-mo hutari-no otoko-ga nutta. every door- QPt two.HUMAN-GEN man-NOM painted 'Two men painted every door. (*scrambled*)'	T1: 6	T1: 39
3.2d	*(Picture as 3.2b)* (*Sentence as 3.2c*)	T1: 19	T1: 26

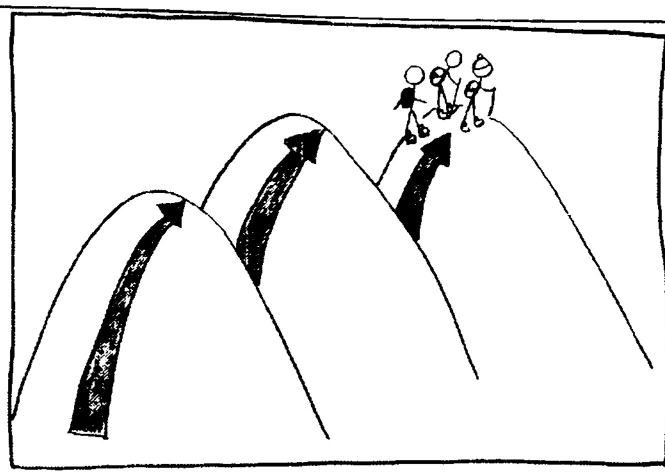
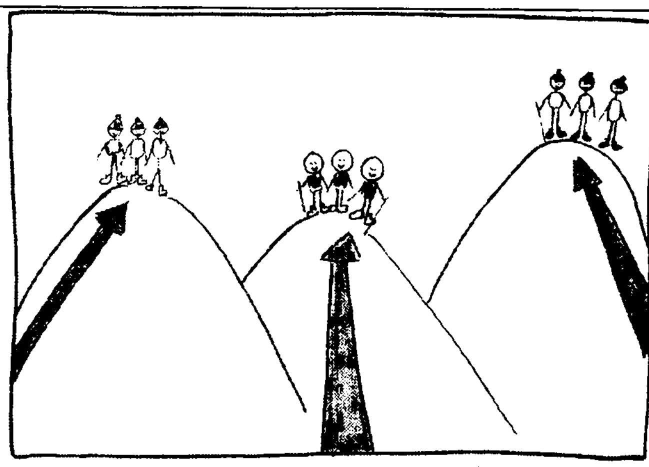
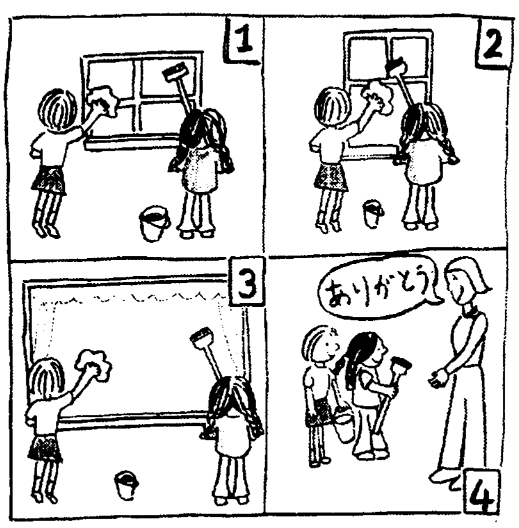
3.3a	 **Translation (following arrows):** Barcelona, Paris, London, Durham, Copenhagen, Berlin, Rome 二人の観光客がどの町も見物した。 Hutari-no kankoukyaku-ga dono mati-mo kenbutu-sita. two.HUMAN-GEN tourist-NOM every town-QPt visited 'Two tourists visited every city.'	T2: 16	T2: 29
3.3b	 **Translation:** Paris, Hong Kong, Honolulu, Kyoto (Sentence as 3.3a)	T2: 22	T2: 23
3.3c	*(Picture as 3.3a)* どの町も二人の観光客が見物した。 Dono mati-mo hutari-no kankoukuaku-ga kenbutu-sita. every door- QPt two.HUMAN-GEN man-NOM painted 'Two tourists visited every city. (*scrambled*)'	T1: 34	T1: 11
3.3d	*(Picture as 3.3b)* (*Sentence as 3.3c*)	T1: 11	T1: 34

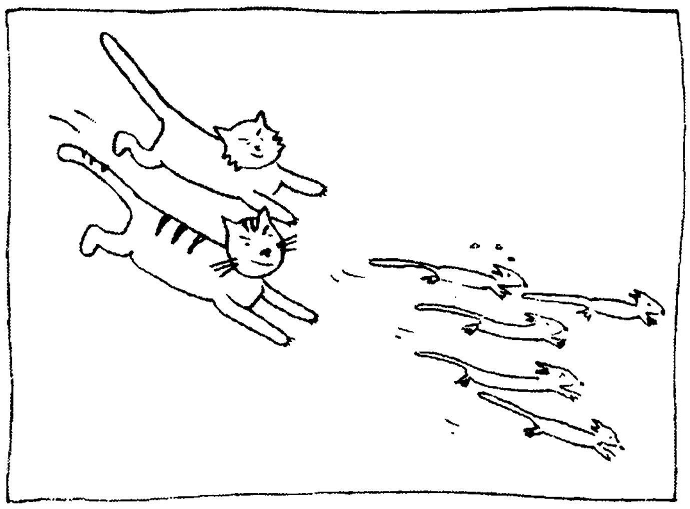
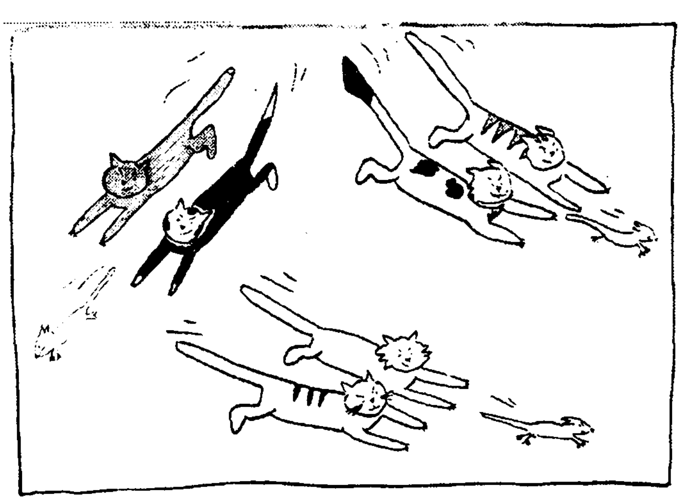
3.4a	 二人の看護婦がどの患者も看病した。 Hutari-no kangohu-ga dono kanzya-mo kanbyou-sita. two.HUMAN-GEN nurse-NOM every patient-QPt looked after 'Two nurses looked after every patient.'	T2: 12	T2: 33
3.4b	 (Sentence as 3.4a)	T2:4	T2: 41
3.4c	*(Picture as 3.4a)* どの患者も二人の看護婦が看病した。 Dono kanzya-mo hutari-no kangohu-ga kanbyou-sita. every patient-QPt two.HUMAN-GEN nurse-NOM looked after 'Two nurses looked after every patient. (*scrambled*)'	T1: 14	T1: 31
3.4d	*(Picture as 3.4b)* (*Sentence as 3.4c*)	T1: 43	T1: 2

3.5a	 三人の男の子がどの太鼓も打った。 Sannin-no otokonoko-ga dono taiko-mo utta. three.HUMAN-GEN boy-NOM every drum-QPt beat 'Three boys beat every drum.'	T2: 23	T2: 22
3.5b	 (Sentence as 3.5a)	T2: 37	T2: 8
3.5c	*(Picture as 3.5a)* どの太鼓も三人の男の子が打った。 Dono taiko-mo sannin-no otokonoko-ga utta. every drum-QP three.HUMAN-GEN boy-NOM beat 'Three boys beat every drum. (*scrambled*)'	T1: 28	T1: 17
3.5d	*(Picture as 3.5b)* (*Sentence as 3.5c*)	T1: 9	T1: 36

4.1a	 三人の男の子がすべての犬を散歩に連れて行った Sannin-no otokonoko-ga subete-no inu-o sanpo ni turete-itta. three.HUMAN-GEN boy-NOM all-GEN dog-ACC walk for took 'Three boys took all the dogs for a walk.'	T1: 3	T1: 42
4.1b	 (Sentence as 4.1a)	T1: 32	T1: 14

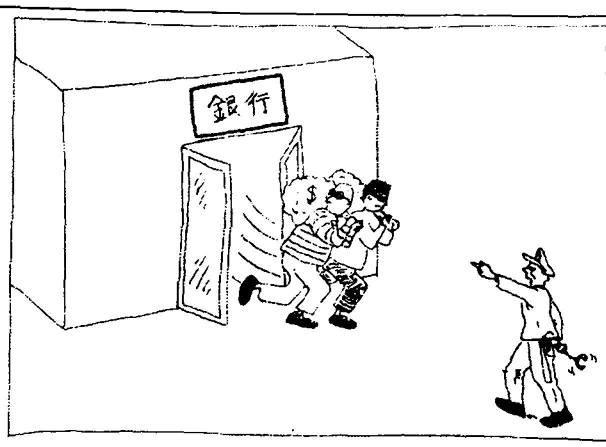
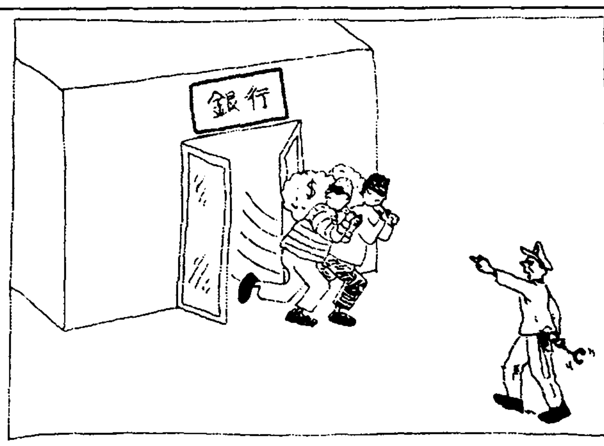
4.2a	 二人の男の子がすべてのテーブルを拭いた。 Hutari-no otokonoko-ga subete-no teeburu-o huita. two.HUMAN-GEN boy-NOM all-GEN table-ACC wiped Two boys wiped all the tables.	T1: 30	T1: 15
4.2b	 (Sentence as 4.2a)	T1: 7	T1: 38

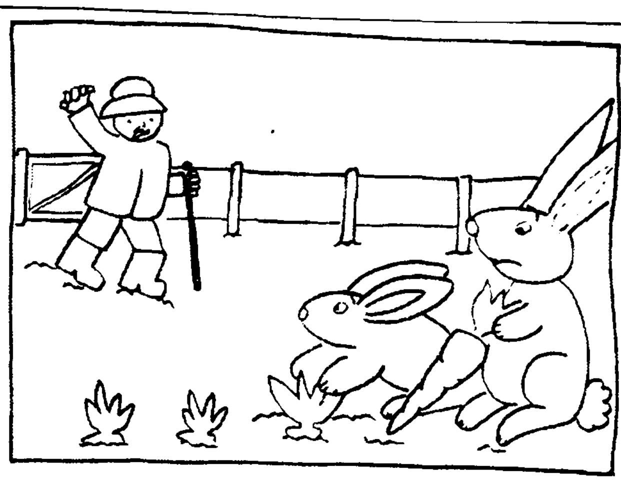
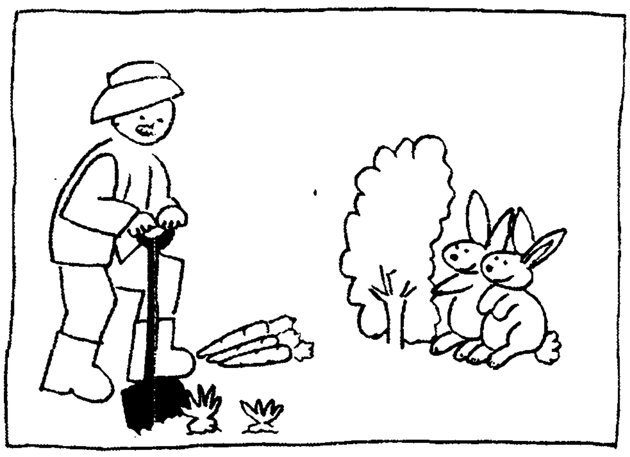
4.3a	 三人の登山客がすべての山に登った。 Sannin-no tozankyaku-ga subete-no yama-ni nobotta. three.HUMAN-GEN hiker-NOM all-GEN mountain-LOC climb 'Three hikers climbed all the mountains.'	T1: 37	T1: 8
4.3b	 (Sentence as 4.3a)	T1: 27	T1: 18
4.4a	 **Translation**: 'Thank you!' 二人の女の子がすべての窓を洗った。 Hutari-no onnanoko-ga subete-no mado-o aratta. two.HUMAN-GEN girl-NOM all-GEN window-ACC washed 'Two girls washed all the windows.'	T1: 20	T1: 25

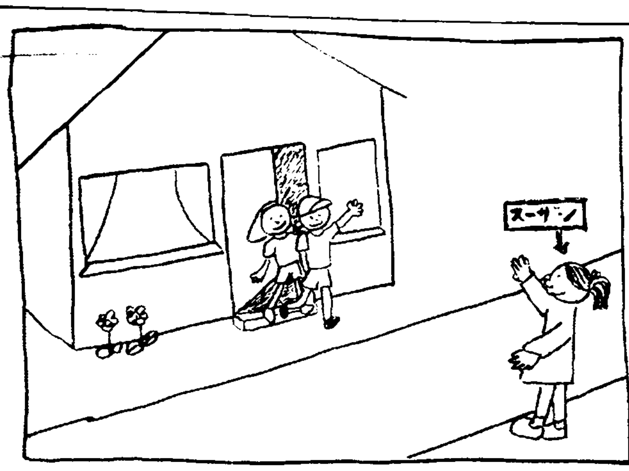
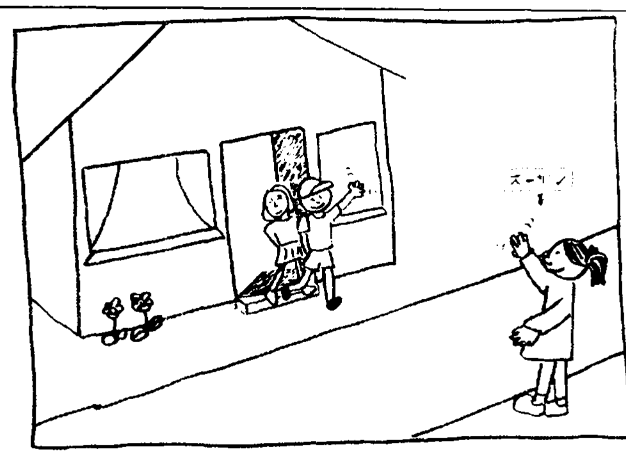
4.4b	 (Sentence as 4.4a)	T1: 41	T1: 4
4.5a	 二匹の猫がすべてのねずみを追った。 Ni-hiki-no neko-ga subete-no nezumi-o otta. two-ANIMAL-GEN cat-NOM all-GEN mouse-ACC chased 'Two cats chased all the mice.'	T1: 18	T1: 27
4.5b	 (Sentence as 4.5a)	T1: 4	T1: 41

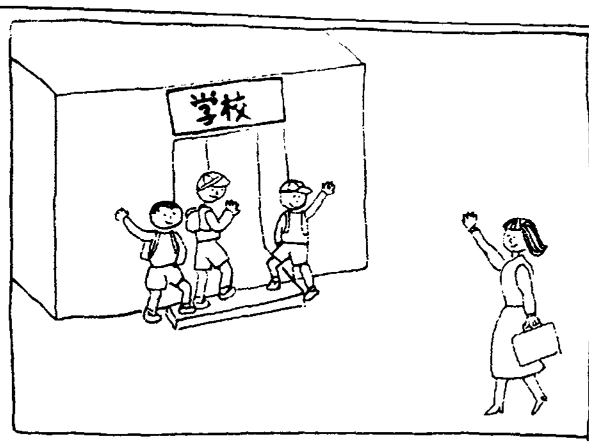
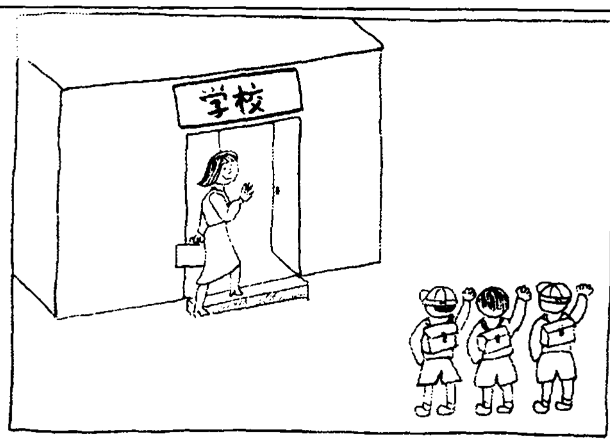
Appendix 5A.ii: Distractor pictures and sentences in Japanese

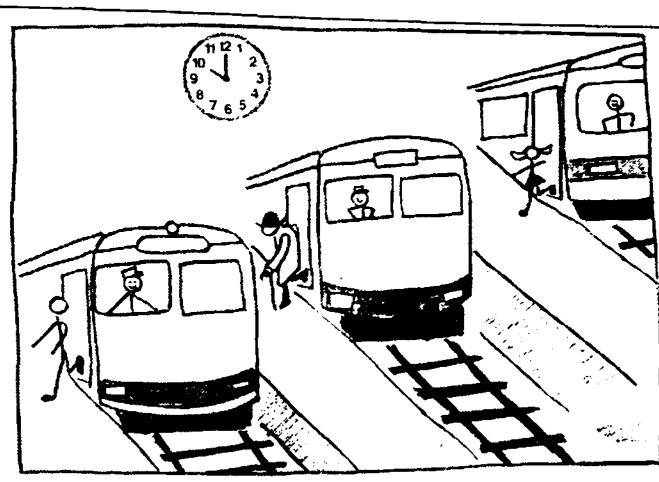
Note: 'T1' and 'T2' in the distractor indices indicate 'QP-QP Task 1' and 'QP-QP Task 2', respectively.

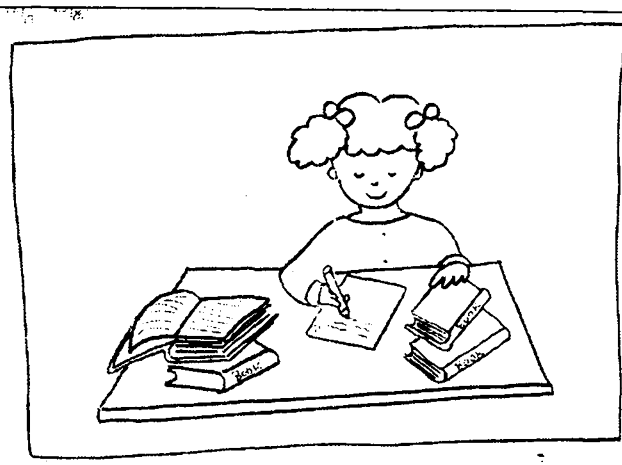
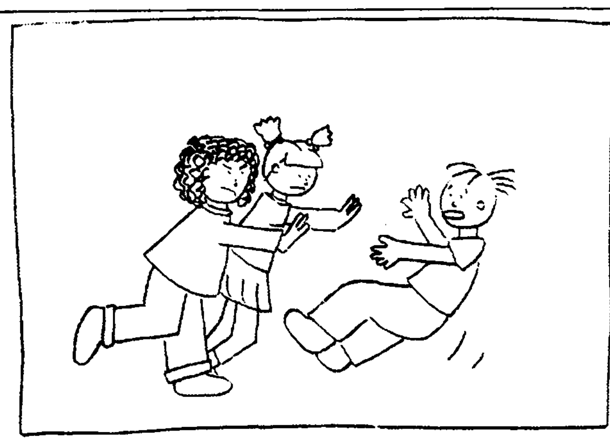
index	picture & sentence	item no.	
		order 1	order 2
T1D01 & T2D01	 **Translation** (above door): Bank		
T1D01	警察官が二人の泥棒を見た。 Keisatukan-ga hutari-no dorobou-o mita. police officer-NOM two.HUMAN-GEN thief-ACC saw 'A policeman saw two thieves.'	T1: 32	T1: 13
T2D01	二人の泥棒を警察官が見た。 Hutari-no dorobou-o keisatukan-ga mita. two.HUMAN- GEN thief-ACC police officer-NOM saw 'A policeman saw two thieves. (*scrambled*)'	T2: 21	T2: 24
T1D02 & T2D02	 **Translation** (above door): Bank		
T1D02	泥棒を警察官が二人とも見た。 Dorobou-o keisatukan-ga hutari tomo mita. thief-ACC police officer-NOM two.HUMAN both saw 'Both policemen saw the thieves. (scrambled)'	T1: 24	T1: 21
T2D02	警察官が二人とも泥棒を見た。 Keisatukan-ga hutari tomo dorobou-o mita. police officer-NOM two.HUMAN both thief-ACC saw 'Both policemen saw the thieves.'	T2: 31	T2: 14

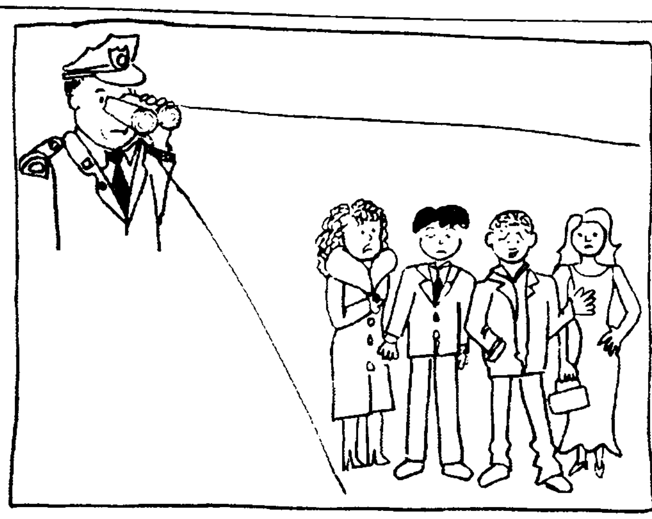
T1D03 & T2D03			
T1D03	だれかが二匹(ひき)のウサギを見た。 Dareka-ga ni-hiki-no usagi-o mita. someone-NOM two-ANIMAL-GEN rabbit-ACC saw 'Someone saw the two rabbits.'	T1: 38	T1: 7
T2D03	二匹(ひき)のウサギをだれかが見た。 ni-hiki-no usagi-o dareka-ga mita. two-ANIMAL-GEN rabbit-ACC someone-NOM saw 'Someone saw the two rabbits. (*scrambled*)'	T2: 14	T2: 31
T1D04 & T2D04			
T1D04	おじさんを二匹(ひき)のウサギが見た。 Ozisan-o ni-hiki-no usagi-ga mita. man-ACC two-ANIMAL-GEN rabbit-NOM saw 'Two rabbits watched the man. (scrambled)'	T1: 44	T1: 1
T2D04	二匹(ひき)のウサギがおじさんを見た。 Ni-hiki-no usagi-ga ozisan-o mita. two-ANIMAL-GEN rabbit-NOM man-ACC saw 'Two rabbits watched the man. (scrambled)'	T2: 38	T2: 7

T1D05 & T2D05	 **Translation:** Susan		
T1D05	子供が二人ともスーザンを見た。 Kodomo-ga hutari tomo Suuzan-o mita. child-NOM two.HUMAN both Susan-ACC saw 'Both children saw Susan.'	T1: 36	T1: 9
T2D05	スーザンを子供が二人とも見た。 Suuzan-o kodomo-ga hutari tomo mita. Susan-ACC child-NOM two.HUMAN both saw 'Both children saw Susan. (*scrambled*)'	T2: 24	T2: 21
T1D06 & T2D06	 **Translation:** Susan		
T1D06	スーザンに子供が二人とも手を振った。 Suuzan ni kodomo-ga hutari tomo te-o hutta. Susan at child-NOM two.HUMAN both hand-ACC waved 'Two children waved at Susan. (*scrambled*)'	T1: 21	T1: 24
T2D06	子供が二人ともスーザンに手を振った。 Kodomo-ga hutari tomo Suuzan ni te-o hutta. child-NOM two.HUMAN both Susan at hand-ACC waved 'Two children waved at Susan.'	T2: 1	T2: 44

T1D07 & T2D07	 **Translation** (above door): School		
T1D07	先生が二人の子供に手を振った。 Sensei-ga hutari-no kodomo ni te-o hutta. teacher-ACC two.HUMAN-GEN child at hand-ACC waved 'The teacher waved at two children.'	T1: 40	T1: 5
T2D07	二人の子供に先生が手を振った。 Hutari-no kodomo ni sensei-ga te-o hutta. two.HUMAN-GEN child at teacher-ACC hand-ACC waved 'The teacher waved at two children. (*scrambled*)'	T2: 9	T2: 36
T1D08 & T2D08	 **Translation** (above door): School		
T1D08	先生に三人の子供が手を振った。 Sensei ni sannin-no kodomo-ga te-o hutta. teacher at three.HUMAN-GEN child-NOM hand-ACC waved 'Three children waved at the teacher. (*scrambled*)'	T1: 33	T1: 12
T2D08	三人の子供が先生に手を振った。 Sannin-no kodomo-ga sensei ni te-o hutta. three.HUMAN-GEN child-NOM teacher at hand-ACC waved 'Three children waved at the teacher.'	T2: 33	T2: 12

T1D09 & T2D09			
T1D09	三人が同じ電車を降りた。 Sannin-ga onazi densya-o orita. three.HUMAN-NOM same train-ACC got off 'Three people got off the same train.'	T1: 15	T1: 30
T2D09	同じホームに三人が降りた。 Onazi hoomu ni sannin-ga orita. same platform on three.HUMAN-NOM got off 'Three people got off on the same platform. (*scrambled*)'	T2: 8	T2: 37
T1D10 & T2D10			
T1D10 & T2D10	皆が同じ車に乗った。 Minna-ga onazyi kuruma ni notta. everyone-NOM same car in got 'Everyone got in the same car.'	T1: 10 T2: 20	T1: 35 T2: 25

T1D11 & T2D11			
T1D11 & T2D11	二人ともコーヒーを飲んだ。 Hutari tomo koohii-o nonda. two.HUMAN both coffee-ACC drank Both people drank coffee.	T1: 1 T2: 44	T1: 44 T2: 1
T1D12 & T2D12			
T1D12 & T2D12	だれかが勉強した。 Dareka-ga benkyou sita. someone-NOM study did Someone studied.	T1: 17 T2: 29	T1: 28 T2: 16
T1D13 & T2D13			
T1D13 & T2D13	男の子を二人の女の子が倒した。 Otokonoko-o hutari-no onnanoko-ga taosita boy-ACC two.HUMAN-GEN girl-NOM pushed over 'Two girls pushed the boy over. (*scrambled*)'	T1: 25 T2: 41	T1: 20 T2: 4

T1D14 & T2D14			
T1D14	五人を警察官（けいさつかん）が見た。 Gonin-o keisatukan-ga mita. five people-ACC police officer-NOM watched 'The detective watched five people. (scrambled)'	T1: 39	T1: 6
T2D14	警察官（けいさつかん）が五人を見た。 Keisatukan-ga gonin-o mita. police officer-NOM five people-ACC watched 'The detective watched five people.'	T2: 15	T2: 30

Appendix 5B: Test sentences in Chinese and English

Notes:

i. See Appendix 5A for the pictures, and for the notes on how to interpret the 'index' and 'item no.' columns.

ii. The English version of the task used the sentences given as translations of the Chinese.

index	sentence	item no.	
		order 1	order 2
1.1a	某人责骂了每个小孩。 mouren zemale meige xiaohai someone scolded every child 'Someone scolded every child.'	T1: 2	T1: 33
1.1b	(*Sentence as 1.1a*)	T1: 21	T1: 14
1.2a	某人尝试了每道菜。 mouren changshile meidao cai someone tried every dish 'Someone tried every dish.'	T1: 7	T1: 23
1.2b	(Sentence as 1.2a)	T1: 9	T1: 26
1.3a	某人抚摸了每只猫 mouren fumole meizhi mao someone stroked every cat 'Someone stroked every cat.'	T1: 19	T1: 16
1.3b	(Sentence as 1.3a)	T1: 16	T1: 19
1.4a	某人读了每本书。 mouren dule meiben shu someone read every book 'Someone read every book.	T1: 10	T1: 25
1.4b	(*Sentence as 1.4a*)	T1: 33	T1: 2
1.5a	某人掷下了每个盘子。 mouren zhixiale meige panzi someone dropped every plate 'Someone dropped every plate.'	T1: 26	T1: 9
1.5b	(*Sentence as 1.5a*)	T1: 5	T1: 30
2.1a	某人打了所有的雨伞。 mouren dale suoyoude yusan someone put up all umbrella 'Someone put up all the umbrellas.'	T2: 4	T2: 31
2.1b	(*Sentence as 2.1a*)	T2: 24	T2: 11
2.2a	某人携带了所有的行李箱。 mouren xiedaile suoyoude xinglixiang someone carried all suitcase 'Someone carried all the suitcases.'	T2: 27	T2: 8

2.2a	(*Sentence as 2.2a*)	T2: 21	T2: 14
2.3a	某人打开了所有的礼物 mouren dakaile suoyoude liwu someone opened all present 'Someone opened all the presents.'	T2: 5	T2: 30
2.3b	(*Sentence as 2.3a*)	T2: 9	T2: 26
2.4a	某人试戴了所有的帽子 mouren shidaile suoyoude maozi someone tried on all hat 'Someone tried on all the hats.'	T2: 28	T2: 7
2.4b	(*Sentence as 2.4a*)	T2: 11	T2: 24
2.5a	某人过了所有的桥 mouren guole suoyoude qiao someone crossed all bridge 'Someone crossed all the bridges.'	T2: 8	T2: 27
2.5b	(*Sentence as 2.5a*)	T2: 20	T2: 15
3.1a	三个女孩放了每只风筝 sange nuhai fangle meizhi fengzheng three girl flew every kite 'Three girls flew every kite.'	T2: 2	T2: 33
3.1b	(*Sentence as 3.1a*)	T2: 33	T2: 2
3.2a	两个男人油漆了每扇门。 liangge nanren youqile meishan men two men painted every door 'Two men painted every door.'	T2: 32	T2: 3
3.2b	(*Sentence as 3.2a*)	T2: 22	T2: 13
3.3a	两个游客游览了每个城市 liangge youke youlanle meige chengshi two tourist visited every city 'Two tourists visited every city.'	T2: 14	T2: 21
3.3b	(*Sentence as 3.3a*)	T2: 17	T2: 28
3.4a	两个护士照顾了每个病人 liangge hushi zhaogule meige bingren two nurses looked after every patient 'Two nurses looked after every patient.'	T2: 3	T2: 32
3.4b	(*Sentence as 3.4a*)	T2: 10	T2: 27
3.5b	三个男孩击打了每个鼓 sange nanhai jidale meige gu three boy beat every drum 'Three boys beat every drum.'	T2: 18	T2: 17
3.5b	(*Sentence as 3.5a*)	T2: 29	T2: 7

4.1a	三个男孩带了所有的狗出去散步 sange nanhai daile suoyoude gou chuqu sanbu three boy took all dog out walk 'Three boys took all the dogs for a walk.'	T1: 3	T1: 32
4.1b	(*Sentence as 4.1a*)	T1: 23	T1: 12
4.2a	两个男孩擦净了所有的桌子. liangge nanhai cajinle suoyoude zhuozi two boy wiped all table Two boys wiped all the tables.	T1: 22	T1: 13
4.2b	(*Sentence as 4.2a*)	T1: 6	T1: 29
4.3a	三个旅行者爬了所有的山。 sange luxingzhe pale suoyoude shan three hiker climbed all mountain 'Three hikers climbed all the mountains.'	T1: 28	T1: 7
4.3b	(*Sentence as 4.3a*)	T1: 20	T1: 15
4.4a	两个女孩清洗了所有的窗户. liangge nuhai qingxile suoyoude chuanghu two girl washed all window 'Two girls washed all the windows.'	T1: 14	T1: 22
4.4b	(*Sentence as 4.4a*)	T1: 32	T1: 3
4.5a	两只猫追捕了所有老鼠. liangzhi mao zhuibule suoyou laoshu two cat chased all mouse 'Two cats chased all the mice.'	T1: 13	T1: 22
4.5b	(*Sentence as 4.5a)*	T1: 4	T1: 31
T1D01	一个警察看见了两个小偷. yige jingcha kanjianle liangge xiaotou a policeman saw two thief 'A policeman saw two thieves.'	T1: 24	T1: 11
T2D01	小偷们被一个警察看到了。 Xiaotoumen bei yige jingcha kandaole. thieves BEI a policeman saw 'The thieves were seen by a policeman.'	T2: 16	T2: 19
T1D02	小偷们被两个警察看到了。 Xiaotoumen bei liangge jingcha kandaole. thieves BEI two policeman saw 'The thieves were seen by both policemen.'	T1: 17	T1: 18
T2D02	两个警察看见了 Liangge jingcha kanjianle xiaotoumen. two policeman saw thieves 'Both policemen saw the thieves.'	T2: 25	T2: 10

T1D03	某人看见了两只兔子。 Mouren kanjianle liangzhi tuzi. someone saw two rabbit 'Someone saw the two rabbits.'	T1: 29	T1: 6
T2D03	两只兔子被某人发现了。 Liangzhi tuzi bei mouren faxianle. two rabbit BEI someone spotted 'The two rabbits were spotted by someone.'	T2: 12	T2: 33
T1D04[214]	男人被两只兔子观看 Nanren bei liangzhi tuzi guankan. man BEI two rabbit watch 'The man was watched by two rabbits.'	T1: 34	T1: 1
T2D04[215]	两个兔子观看了那个男人。 Liangge tuzi guankanle nage nanren. two rabbit watched that man 'Two rabbits watched the man.'	T2: 30	T2: 5
T1D05	两个小孩都看见了苏珊。 Liangge xiaohai dou kanjianle Susan. two child all saw Susan 'Both children saw Susan.'	T1: 27	T1: 8
T2D05	两个小孩看到了苏珊。 Liangge xiaohai kandaole Susan two child saw Susan 'Two children saw Susan.'	T2: 19	T2: 16
T1D06[216]	两个小孩向苏珊挥手。 Liangge xiaohai xiang Susan huishou. two child at Susan wave 'Two children waved at Susan.'	T1: 15	T1: 20

[214] Distractors T1D04 and T2D04 were excluded as unreliable in Chinese because they were neither clearly accepted nor clearly rejected (see Chapter 5, Section 5.2.3.1). They were intended to be acceptable. The following factors may have contributed to some individuals nonetheless rejecting them. First, the verb *guankan* 'watch' may have seemed unnatural with 'rabbits' as its subject, because *guankan* 'watch' implies conscious thought, which may be incompatible with a non-human subject. Second, the passive form (T1D04 only) in Chinese is also rarely used with a non-human agent. (Thanks to Yu Jiang, Li-Chen Li and Zhengzheng Wang for discussion, April 2004.)

[215] See footnote 214.

[216] Distractor T1D06 was unreliable in both Chinese and English, and distractor T2D06 was unreliable in Chinese (see Chapter 5 Section 5.2.3.1). The picture was intended to depict only one child waving, the other child with her hand behind her back (see Appendix 5A). Thus, the items were intended to be unacceptable. The lack of consensus in Chinese and English. may be due to some individuals noting that an act of waving was depicted, and that it was associated with two children (despite only one having a hand in the air). This may have been enough to allow some participants to accept the sentence. In native Japanese and Korean, these distractors were rejected, as expected. However, in Japanese and Korean the verbal expression indicating 'waved' is more precise than in English and Chinese, since it includes the word for 'hand': *te-o hutta* (Japanese)/*son-ul huntulassta* (Korean) 'hand-ACC waved'. Inclusion of the word 'hand' in Japanese and Korean may have drawn attention to the children's hands in the picture and led to uniform acknowledgement of a picture-sentence mismatch.

T2D06[217]	两个小孩都向苏珊挥手。 Liangge xiaohai dou xiang Susan huishou. two child all at Susan wave 'Both children waved at Susan.'	T2: 1	T2: 34
T1D07	老师向两个小孩挥手。 Laoshi xiang liangge xiaohai huishou. teacher at two child wave 'The teacher waved at both children.'	T1: 31	T1: 4
T2D07	教师向孩子们挥了手。 Jiaoshi xiang haizimen huileshou. teacher at children wave 'The teachers waved at the children.'	T2: 7	T2: 28
T1D08	三个小孩向老师挥手。 Sange xiaohai xiang laoshi huishou. three child at teacher wave 'Three children waved at the teacher.'	T1: 25	T1: 10
T2D08	三个小孩向老师挥手。 Sange xiaohai xiang laoshi huishou. three child at teacher wave 'Three children waved at the teacher.'	T2: 26	T2: 9
T1D09	三个人下了同一列火车。 Sange ren xiale tongyilie huoche. three people got off same train 'Three people got off the same train.'	T1: 11	T1: 24
T2D09	三个人从同一个站台下了车。 Sangeren cong tongyige zhantai xialeche. three people from samedi platform got off 'Three people got off on the same platform.'	T2: 11	T2: 24
T1D10 & T2D10	每个人都上了同一辆汽车。 Meigeren dou shangle tongyiliang qiche. everyone all got.in same car 'Everyone got in the same car.'	T1: 8 T2: 15	T1: 27 T2: 20
T1D11 & T2D11	两个人都喝了咖啡。 Lianggeren dou hele kafei. two people all drank coffee 'Both people drank coffee.'	T1: 1 T2: 34	T1: 34 T2: 1
T1D12 & T2D12	某人学习了。 Mouren xuexile. someone studied 'Someone studied.'	T1: 17 T2: 23	T1: 33 T2: 12
T1D13 & T2D13	男孩被两个女孩推倒了。 Nanhai bei liangge nuhai tuidaole. boy BEI two girl pushed down The boy was pushed over by two girls.	T1: 18 T2: 31	T1: 17 T2: 4

[217] See footnote 216.

T1D14	这个警察看到了五个人。 Zhege jingcha kandaole wugeren. this policeman saw five people 'The detective watched five people.'	T1: 30	T1: 5
T2D14	四个人被一个警察观察过。 Sigeren bei yige jingcha guanchaguo. four.people BEI a policeman watched Four people were watched by the detective.	T2: 13	T2: 22

Appendix 5C: Test sentences in Korean

Notes:

i. See Appendix 5A for the pictures, and for the notes on how to interpret the 'index' and 'item no.' columns.

ii. Korean verbal suffixes, such as *–ss* (past tense morpheme), *-ko* (complementiser), and *–ta* (declarative morpheme), and the post-proper-name phonetic filler *i* are not glossed separately.

index	sentence	item no.
K1.1a	누군가가 모든 아이를꾸짖었다. Nwukwunka-ka motun ai-lul kkwucicessta someone-NOM all child-ACC scolded 'Someone scolded all the children.'	T1: 2
K1.1b	(*Sentence as K1.1a*)	T1: 23
K1.1c	모든 아이를 누군가가 꾸짖었다. Motun ai-lul nwukwunka-ka kkwucicessta all child-ACC someone-NOM scolded 'Someone scolded all the children. (*scrambled*)'	T2: 29
K1.1d	(*Sentence as K1.1c*)	T2: 13
K1.2a	누군가가 모든 요리를 맛보았다. Nwukwunka-ka motun yoli-lul maspoassta someone-NOM all dish-ACC tried 'Someone tried all the dishes.'	T1: 5
K1.2b	(Sentence as K1.2a)	T1: 9
K1.2c	모든 요리를 누군가가 맛보았다. Motun yoli-lul nwukwunka-ka maspoassta All dish-ACC someone-NOM tried 'Someone tried all the dishes. (*scrambled*)'	T2: 5
K1.2d	(*Sentence as K1.2c*)	T2: 12
K1.3a	누군가가 모든 고양이를 쓰다듬었다. Nwukwunka-ka motun koyangi-lul ssutatumessta someone-NOM all cat-ACC stroked 'Someone stroked all the cats.'	T1: 21
K1.3b	(*Sentence as K1.3a*)	T1: 18
K1.3c	모든 고양이를 누군가가 쓰다듬었다. Motun koyangi-lul nwukwunka-ka ssutatumessta all cat-ACC someone-NOM stroked 'Someone stroked all the cats. (*scrambled*)'	T2: 20
K1.3d	(*Sentence as K1.3c*)	T2: 30
K1.4a	누군가가 모든 책을 읽었다. Nwukwunka-ka motun chayk-ul ilkessta. someone-NOM all book-ACC read 'Someone read all the books.	T1: 10

K1.4b	(*Sentence as K1.4a*)	T1: 32
K1.4c	모든 책을 누군가가 읽었다. Motun chayk-ul nwukwunka-ka ilkessta. all book-ACC someone-NOM read 'Someone read all the books. (*scrambled*)'	T2: 3
K1.4d	(*Sentence as K1.4c*)	T2: 26
K1.5a	누군가가 모든 접시를 떨어뜨렸다. Nwukwunka-ka motun cepsi-lul ttelettulyessta. someone-NOM all plate-ACC dropped 'Someone dropped all the plates.'	T1: 27
K1.5a	(*Sentence as K1.5a*)	T1: 3
K1.5c	모든 접시를 누군가가 떨어뜨렸다. Motun cepsi-lul nwukwunka-ka ttelettulyessta. all plate-ACC someone-NOM dropped 'Someone dropped all the plates. (*scrambled*)'	T2: 14
K1.5d	(*Sentence as K1.5c*)	T2: 24
K2.1a	세명의 소녀가 모든 연을 날렸다. Sey-myeng-uy sonye-ka motun yen-ul nallyessta. three-CL.HUMAN-GEN girl-NOM all kite-ACC flew 'Three girls flew all the kites.'	T2: 2
K2.1b	(*Sentence as K2.1a*)	T2: 33
K2.1c	모든 연을 세명의 소녀가 날렸다. Motun yen-ul sey-myeng-uy sonye-ka nallyessta. all kite-ACC three-CL.HUMAN-NOM flew 'Three girls flew all the kites. (*scrambled*)'	T1: 13
K2.1d	(*Sentence as K2.1c*)	T1: 17
K2.2a	두명의 남자가 모든 문을 칠했다 Twu-myeng-uy namca-ka motun mwun-ul chilhayssta. three-CL.HUMAN-GEN man-NOM all door-ACC painted 'Two men painted all the doors.'	T2: 32
K2.2a	(*Sentence as K2.2a*)	T2: 21
K2.2c	모든 문을 두명의 남자가 칠했다. Motun mwun-ul twu-myeng-uy namca-ka chilhayssta. all door-ACC three-CL.HUMAN-GEN man-NOM painted 'Two men painted all the doors. (*scrambled*)'	T1: 4
K2.2d	(*Sentence as K2.2c*)	T1: 15
K2.3a	두명의 여행객이 모든 도시를 방문했다. Twu-myeng-uy yehayngkayk-i motun tosi-lul pangmwunhayssta. two-CL.HUMAN-GEN tourist-NOM all city-ACC visited 'Two tourists visited all the cities.'	T2: 11

K2.3b	(*Sentence as K2.3a*)	T2: 17
K2.3c	모든 도시를 두명의 여행객이 방문했다. Motun tosi-lul twu-myeng-uy yehayngkayk-i pangmwunhayssta. all city-ACC two-CL.HUMAN-GEN tourist-NOM visited 'Two tourists visited all the cities. (*scrambled*)'	T1: 26
K2.3d	(*Sentence as K2.3c*)	T1: 8
K2.4a	두명의 간호원이 모든 환자를 간호했다. Twu-myeng-uy kanhowen-i motun hwanca-lul kanhohayssta. two-CL.HUMAN-GEN nurse-NOM all patient-ACC looked after 'Two nurses looked after all the patients.'	T2: 4
K2.4b	(*Sentence as K2.4a*)	T2: 8
K2.4c	모든 환자를 두명의 간호원이 간호했다. Motun hwanca-lul twu-myeng-uy kanhowen-i kanhohayssta. all patient-ACC two-CL.HUMAN-GEN nurse-NOM looked after 'Two nurses looked after all the patients. (*scrambled*)'	T1: 11
K2.4d	(*Sentence as K2.4c*)	T1: 33
K2.5a	세명의 소년이 모든 북을 쳤다. Sey-myeng-uy sonyen-i motun pwuk-ul chyessta. three-CL.HUMAN-GEN boy-NOM all drum-ACC beat 'Three boys beat all the drums.'	T2: 18
K2.5b	(*Sentence as K2.5a*)	T2: 27
K2.5c	모든 북을 세명의 소년이 쳤다. motun pwuk-ul sey-myeng-uy sonyen-i chyessta. all drum-ACC three-CL.HUMAN-GEN boy-NOM beat 'Three boys beat all the drums. (*scrambled*)'	T1: 22
K2.5d	(*Sentence as K2.5c*)	T1: 6
T1D01	경찰관이 두 도둑을 보았다. Kyengchalkwan-i twu totwuk-ul poassta. policeman-NOM two thief-ACC saw 'A policeman saw two thieves.'	T1: 24
T2D01	두명의 도둑을 경찰관이 보았다. Twu-myeng-uy totwuk-ul kyengchalkwan-i poassta. two CL.HUMAN-GEN thief-ACC policeman-NOM saw 'A policeman saw two thieves. (scrambled)'	T2: 16
T1D02	도둑을 두 경찰관이 봤다. Totwuk-ul twu kyengchalkwan-i pwassta. thief-ACC two policeman-NOM saw 'Two policemen saw the thieves. (scrambled)'	T1: 19

T2D02	두명의 경찰관이도 도둑들을 보았다. Twu-myeng-uy kyengchalkwan-i-to totwuk-tul-ul poassta. two CL.HUMAN-GEN policeman-NOM-both thief-PL-ACC saw 'Both policemen saw the thieves.'	T2: 23
T1D03	누군가가 두 마리의 토끼를 보았다. Nwukwunka-ka twu-mori-uy thokki-lul poassta. someone-NOM two-CL-GEN rabbit-ACC saw 'Someone saw two rabbits.'	T1: 29
T2D03	두 마리의 토끼를 누군가가 발견했다 Twu-mori-uy thokki-lul nwukwunka-ka poassta. two-CL-GEN rabbit-ACC someone-NOM saw 'Someone saw two rabbits.'(scrambled)'	T2: 9
T1D04[218]	그 사람이 두마리의 토끼를 보았다. Ku saram-i twu-mori-uy thokki-lul poassta. that person-NOM two-CL-GEN rabbit-ACC saw 'The man saw two rabbits.'	T1: 34
T2D04[219]	두 마리의 토끼를 그 남자가 보았다. Twu-mori-uy thokki-lul ku namca-i poassta. two-CL-GEN rabbit-ACC that man-NOM saw 'Two rabbits watched the man. (scrambled)'	T2: 10
T1D05	두 아이가 순이를 보았다. Twu ai-ka Swun-i-lul poassta. two child-NOM Swun-ACC saw 'Both children saw Susan.'	T1: 28
T2D05	영이를 두명의 아이가 보았다. Yang-i-lul twu-myeng-uy ai-ka poassta. Yang-ACC two CL.HUMAN-GEN child-NOM saw 'Two children saw Yang. (*scrambled*)'	T2: 19
T1D06 & T2D06	두 아이가 순이에게 손을 흔들었다. Twu ai-ka Swun-i-eykey son-ul huntulassta. two child-NOM Swun-at hand-ACC waved 'Two children waved at Susan.'	T1: 16 T2: 1
T1D07	선생님이 두 아이에게 손을 흔들었다. Sensayngnim-i twu ai eykey son-ul huntulessta. teacher-NOM two child at hand-ACC waved 'The teacher waved at two children.'	T1: 31

[218] This distractor was excluded as unreliable in Korean because it was not clearly accepted or rejected (see Chapter 5, Section 5.2.3.1). This was probably due to the sentence actually being a mistranslation. It should have been 'Two rabbits watched the man' with SOV word order. In the picture (see Appendix 5), the man is not clearly looking towards the rabbits. This is likely to have been the source of confusion on this item. (One native Korean respondent commented that this picture had been confusing due to not knowing where the man was looking.)
[219] This distractor was also excluded as unreliable in Korean due to not being clearly accepted or rejected (see Chapter 5, Section 5.2.3.1). The reasons for the lack of concensus about this item are unclear.

T2D07	두 아이에게 선생님이 손을 흔들었다. Twu ai eykey sensayngnim-i son-ul huntulessta. two child at teacher-NOM hand-ACC waved 'The teacher waved at two children. (*scrambled*)'	T2: 7
T1D08	선생님에게 세명의 아이들이 손을 흔들었다. sensayngnim eykey sey-myeng-uy ai-tul-i son-ul huntulessta. teacher at three-CL.HUMAN-GEN child-PL-NOM hand-ACC waved 'Three children waved at the teacher. (*scrambled*)'	T1: 25
T2D08	세명의 아이가 선생님에게 손을 흔들었다. sey-myeng-uy ai-ka sensayngnim eykey son-ul huntulessta. three-CL.HUMAN-GEN child-NOM teacher at hand-ACC waved 'Three children waved at the teacher.'	T2: 25
T1D09	세 사람이 같은 기차에서 내렸다. Sey saram-i kalun kicha eyse naylyessta. three person-NOM same train from alighted 'Three people got off the same train.'	T1: 12
T2D09	세명의 사람이 같은 승강구에 내렸다. sey-myeng-uy saram-i kalun sungkangkwu ey naylyessta. three-CL.HUMAN-GEN person-NOM same platform to alighted 'Three people got off on the same platform.'	T2: 6
T1D10 & T2D10	누구나 같은 자동차를 탔다. Nwukwuna kalun catongcha-lul thassta. Everyone same car-ACC got in 'Everyone got in the same car.'	T1: 7 T2: 15
T1D11 & T2D11	두 사람이 커피를 마셨다. Twu saram-i khephi-lul masyessta. two person-NOM coffee-ACC drank 'Two people drank coffee.'	T1: 1 T2: 34
T1D12 & T2D12	누군가가 공부했다. Nwukwunka-ka kongpwuhayssta. someone-NOM studied 'Someone studied.'	T1: 14 T2: 31
T1D13 & T2D13	그 소년을 두 소녀가 떼밀었다. Ku sonyen-ul twu sonye-ka tteymilassta. the boy-ACC two girl-NOM pushed over 'Two girls pushed the boy over. (*scrambled*)'	T1: 20 T2: 10
T1D14	다섯 사람을 경찰관이 봤다. Tases saram-ul kyengchalkwan-i pwassta. five person-ACC detective-NOM watched 'The detective watched five people. (scrambled)'	T1: 30
T2D14	네명의 사람이 탐정에 의해 감시당했다. Ney-myeng-uy saram-i thamceng ey-uyhay kamsitanghayssta. four-CL.HUMAN.GEN person-NOM detective by watched 'Four people were watched by the detective.'	T2: 10

Appendix 5D: L2ers' answer sheet in English

Personal Details

1. Your initials: .... 2. Your age:.................. 3. Gender: M F
4. What is (are) your native language(s)? ....
   If you have more than one native language, which language is dominant, if any? Please give details:...
   ....
5. Please describe your formal study of Japanese by filling in this table:

Type of instruction	When
Example: *degree in Japanese at UK university, approx. 8 hrs per week of instruction*	*Oct 2001-present*

6. Have you ever lived in Japan? yes no
   If yes, how long did you live there? ....

1

Instructions (Task 1)

For each test item you will see a picture. Underneath the picture, a sentence will be shown. Please judge whether the sentence is a possible description of the picture. In other words, does the sentence match the picture? Indicate your answer by circling one of the options on the scale on your answersheet. The scale is as follows:

No, definitely not			Yes, perfectly		Can't decide
-2	-1	+1	+2		X

NOTE:

All the Japanese sentences are grammatically correct. This is not a 'find the mistakes' task.

Examples

(Does the sentence match the picture?)

	No, definitely not			Yes, perfectly		Can't decide
Ex. 1	-2	-1	+1	+2		X
Ex. 2	-2	-1	+1	+2		X
Ex. 3	-2	-1	+1	+2		X
Ex. 4	-2	-1	+1	+2		X
Ex. 5	-2	-1	+1	+2		X
Ex. 6	-2	-1	+1	+2		X

****

Post-task comments (Please circle your answer, where appropriate)

1. Was the time for each test item: `too long` `just right` `too short`
2. Did you find any of the pictures difficult to understand?

   ....

   ....
3. Any further comments?

   ....

   ....

2

	No, definitely not			Yes, perfectly
1	−2	−1	+1	+2
2	−2	−1	+1	+2
3	−2	−1	+1	+2
4	−2	−1	+1	+2
5	−2	−1	+1	+2
6	−2	−1	+1	+2
7	−2	−1	+1	+2
8	−2	−1	+1	+2
9	−2	−1	+1	+2
10	−2	−1	+1	+2
11	−2	−1	+1	+2
12	−2	−1	+1	+2
13	−2	−1	+1	+2
14	−2	−1	+1	+2
15	−2	−1	+1	+2
16	−2	−1	+1	+2
17	−2	−1	+1	+2
18	−2	−1	+1	+2
19	−2	−1	+1	+2
20	−2	−1	+1	+2
21	−2	−1	+1	+2
22	−2	−1	+1	+2
23	−2	−1	+1	+2
24	−2	−1	+1	+2

Can't decide
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X

3

25	–2	–1	+1	+2	X
26	–2	–1	+1	+2	X
27	–2	–1	+1	+2	X
28	–2	–1	+1	+2	X
29	–2	–1	+1	+2	X
30	–2	–1	+1	+2	X
31	–2	–1	+1	+2	X
32	–2	–1	+1	+2	X
33	–2	–1	+1	+2	X
34	–2	–1	+1	+2	X
35	–2	–1	+1	+2	X
36	–2	–1	+1	+2	X
37	–2	–1	+1	+2	X
38	–2	–1	+1	+2	X
39	–2	–1	+1	+2	X
40	–2	–1	+1	+2	X
41	–2	–1	+1	+2	X
42	–2	–1	+1	+2	X
43	–2	–1	+1	+2	X
44	–2	–1	+1	+2	X

4

Appendix 5E: L2ers' answer sheet in Japanese

Personal Details

1. イニシャル: .... 2. 年令:.................. 3. 性別: 男 女

4. 第一言語は何ですか。 ....

(中国語の場合、どの方言か書いて下さい：..................)

子供のころからバイリンガルの場合どの言語を主に使いますか。..................

....

5. 現在までの日本語学習について書いて下さい:

学習形態	いつからいつまで
例：高校で、週に3時間 大学で、週に6時間 日本の大学で、留学生として、週に10時間	4月1997年~2月2000年 4月2000年~2月2002年 4月2002年~現在

6. いつから日本に住んでいますか。 ....

7. 日本語以外の外国語を勉強したことがありますか。 はい いいえ

「はい」の場合、その学習について書いて下さい

学習形態	いつからいつまで

1

テストの方法（タスク１）
このテストはみなさんに絵と文の組み合わせを見ていただいて、各組の絵と文が合っているかどうか判断していただくものです。判断は用紙のスケールを使い、当てはまる数字を○で囲んで下さい。スケールは次の通りです：

（文は絵に合っていますか。）

いいえ、少しも合っていない			はい、完全に合っている	分からない
−2	−1	+1	+2	X

注：テストの文はどれも文法的に正しいので、これは「間違いを探す」ようなテストではありません。

例：
（文は絵に合っていますか。）

	いいえ、少しも合っていない			はい、完全に合っている	分からない
Ex. 1	−2	−1	+1	+2	X
Ex. 2	−2	−1	+1	+2	X
Ex. 3	−2	−1	+1	+2	X
Ex. 4	−2	−1	+1	+2	X
Ex. 5	−2	−1	+1	+2	X
Ex. 6	−2	−1	+1	+2	X

****

テスト後のコメント

1. **各件の時間は：　　長過ぎた　　ちょうどよかった　　短過ぎた**
2. **分かりにくい絵がありましたか。（どれか覚えていますか。）**

   ……………………………………………………………………

   ……………………………………………………………………

3. **その他のコメントはありますか。**

   ……………………………………………………………………

   ……………………………………………………………………

2

（文は絵に合っていますか。）

	いいえ、少しも合っていない			はい、完全に合っている	分からない
1	–2	–1	+1	+2	X
2	–2	–1	+1	+2	X
3	–2	–1	+1	+2	X
4	–2	–1	+1	+2	X
5	–2	–1	+1	+2	X
6	–2	–1	+1	+2	X
7	–2	–1	+1	+2	X
8	–2	–1	+1	+2	X
9	–2	–1	+1	+2	X
10	–2	–1	+1	+2	X
11	–2	–1	+1	+2	X
12	–2	–1	+1	+2	X
13	–2	–1	+1	+2	X
14	–2	–1	+1	+2	X
15	–2	–1	+1	+2	X
16	–2	–1	+1	+2	X
17	–2	–1	+1	+2	X
18	–2	–1	+1	+2	X
19	–2	–1	+1	+2	X
20	–2	–1	+1	+2	X
21	–2	–1	+1	+2	X
22	–2	–1	+1	+2	X
23	–2	–1	+1	+2	X
24	–2	–1	+1	+2	X

3

25	–2	–1	+1	+2
26	–2	–1	+1	+2
27	–2	–1	+1	+2
28	–2	–1	+1	+2
29	–2	–1	+1	+2
30	–2	–1	+1	+2
31	–2	–1	+1	+2
32	–2	–1	+1	+2
33	–2	–1	+1	+2
34	–2	–1	+1	+2
35	–2	–1	+1	+2
36	–2	–1	+1	+2
37	–2	–1	+1	+2
38	–2	–1	+1	+2
39	–2	–1	+1	+2
40	–2	–1	+1	+2
41	–2	–1	+1	+2
42	–2	–1	+1	+2
43	–2	–1	+1	+2
44	–2	–1	+1	+2

X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X

4

Appendix 5F: Tables of results

5F.1. Raw data

Table 5F.1.i: Mean ratings on Japanese, English and Chinese test types by participant[220]

	id	order	mean rating (scale 0 to 3)											
			1a	1b	1c	1d	2a	2b	3a	3b	3c	3d	4a	4b
JJ	JJ19	1	2	0.2	1.6	2.75	2	0	2.4	0.2	1.75	1.75	2.2	0.4
	JJ20	1	2.8	0	2.8	2.4	2.8	0.8	2.8	1.8	2.8	3	2.2	1.6
	JJ22	1	2.8	0.8	2.4	1.4	3	0.8	3	1	2.2	2.6	3	0.8
	JJ23	1	1.2	0.2	1.2	2.4	1	0.8	3	0.4	3	2.8	3	0.2
	JJ24	1	1.4	0.8	1.2	1.6	2.8	0.4	2.8	0.4	2.8	1.8	3	0.4
	JJ25	1	2.2	0.8	1.6	1.4	2.4	1	2	1.8	2.2	2.4	2.6	1.6
	JJ26	1	2.8	0.4	2.6	2	2.8	0.2	3	1	3	3	2.8	1.6
	JJ27	1	2	1	3	2.8	3	0.4	3	0	2.8	2	3	0.2
	JJ28	1	2.6	1.6	2.8	2.4	2.8	1.8	3	1.2	2.6	1.6	2.6	1.4
	JJ29	1	3	0.6	3	2.6	3	0.6	3	0.6	2.8	2.8	2.6	0.4
	JJ30	1	2.6	0.4	1.8	2.6	2.2	0.6	3	1.8	2.8	2.6	2.6	1
	JJ31	1	2.4	0.2	1.4	2.4	2.4	0.6	2.6	0.4	2.4	2	2.4	0.2
	JJ32	1	2.4	0	3	1	3	0	3	0	2.6	1.8	3	0
	JJ33	3	2.2	1	2.6	2	3	0.8	3	0	2.75	3	3	0
	JJ34	3	2.2	0	1.4	3	2.2	0.2	2.6	0.2	2.8	3	3	0.6
	JJ35	3	1.8	1.2	1.2	2.6	1.6	0	3	0.4	2	1.8	2.6	0.2
	JJ36	3	3	3	1.2	2.4	2.8	0.4	3	0.8	3	3	2	2.4
	JJ38	3	2.8	1.2	2.4	2.8	3	0.2	3	0.6	3	2.2	2.6	0.6
	JJ39	3	0.2	0.2	1.2	2.2	2.2	0.4	3	0	2.2	3	2.8	0.6
	JJ40	3	2.8	0.2	1.8	3	2.8	2	3	1.4	2.6	3	3	0.4
EJ int	EJ13	1	2.2	1.4	2.8	1.6	2.6	1.2	2.8	1	2.4	1.4	2.8	1
	EJ28	2	2.6	1.25	2.2	2.2	3	1.8	2.4	1.4	1.8	1.6	2	1.2
	EJ30	2	2.4	2	2.8	2.2	2.6	1.8	2.6	1.8	2.2	2.6	2.8	2.2
	EJ17	2	3	1	2.8	2	3	1.4	3	1.4	2.2	1.8	3	1
	EJ03	1	2.75	2.4	3	2.8	3	2.2	3	2.6	3	2.8	3	2.8
	EJ16	2	3	2.8	2	3	3	1.8	3	3	3	3	3	2.2
	EJ19	2	3	1.4	2.8	2.6	3	2	2.8	2.2	2.6	1.8	3	1.2
	EJ25	2	3	0.4	2.4	1.4	2.8	0.4	3	1	2.2	1.2	3	0.6
	EJ02	1	2.8	1.6	3	3	3	2.2	3	2.6	2.6	1.8	2.8	0.6
	EJ08	1	3	1.6	3	3	3	3	3	3	2.6	1.8	3	1.4
	EJ09	1	2.6	1.6	2.8	3	2.8	2.6	3	2.8	2	1.6	2.8	1
	EJ14	1	2.4	1.8	2.8	2.8	2.8	2	2.8	0.8	1.6	1.4	2.6	0.4
	EJ26	2	2.2	2.2	2.25	2.8	2.6	2.2	2	2.8	2.75	2.8	2.33	2.6
	EJ27	2	2.8	1.6	2	3	2.8	1	2.8	1	2.2	2.2	2.8	1.2
	EJ29	2	2.8	0.6	3	1.8	3	0.8	3	1.6	3	1.6	3	0.6
	EJ22	2	3	1	2.6	2.2	2.8	1.4	2.6	2	2.5	2	2.4	1.4
	EJ11	1	2.8	2	2.4	2.2	3	2	2.8	2.4	2.6	2	3	1.6
	EJ07	1	3	2.4	3	3	3	3	3	1.2	2.8	1.8	3	0.8
EJ adv	EJ05	1	3	0.6	3	0.4	3	0	3	0	2.8	1.8	2.8	0.4
	EJ23	2	3	0.8	3	0.8	3	0.8	2.8	1	2.2	1	2.8	0.8
	EJ24	2	2.8	1.2	2	1.6	3	1	3	1.4	3	1.4	3	1
	EJ06	1	3	0.6	3	2.2	3	0.8	3	0.8	3	2	3	1
	EJ12	1	3	0.8	3	1	3	1	3	1.2	2.6	2.2	3	1
	EJ04	1	3	0	3	0	3	0	3	0	2.4	0.6	3	0
	EJ20	2	2.8	2.6	3	2.8	3	3	3	2.4	2.8	3	3	1
	EJ15	2	1.8	2.8	1.2	3	2.8	1	2.6	2	3	2.6	3	1.2

[220] See Appendix 3 for proficiency scores for learners and for age and gender of all participants.

	id	order	mean rating (scale 0 to 3)											
			1a	1b	1c	1d	2a	2b	3a	3b	3c	3d	4a	4b
CJ int	CJ10	2	0.8	1.6	1.2	2.2	0.8	0.2	1.6	1.6	1.2	2	0.8	1
	CJ12	2	2	2.4	2	2	1.2	2	0.8	1.2	2.6	2.2	1.2	1.6
	CJ11	2	1	0.8	0.8	1.2	1.4	0.4	2.4	1.4	1.4	2.2	1.4	0.6
	CJ13	2	3	0.75	3	0	3	0	3	0	3	3	3	0
	CJ01	2	3	1.4	3	3	3	0	3	0.6	3	2.6	3	0
	CJ04	2	2.6	0.2	2.2	2.6	3	0.8	2.2	0.6	2.4	1.8	2.6	0.6
CJ adv	CJ03	2	2.8	1.8	3	2.2	3	0.6	3	0.4	2.6	2	2.8	0.6
	CJ15	2	3	1.2	3	2.6	3	0.4	2.8	0.8	2.8	2	3	1.6
	CJ09	1	2.6	0.8	2	2.4	2.4	1.2	2.4	0.8	2.4	2.6	2.4	1
	CJ18	1	3	0.2	2.8	2.4	3	0	3	0.2	3	2	3	0.2
	CJ16	1	2.6	1.8	2.4	2	2	1	2.8	1.4	2.8	2.4	3	1.6
	CJ07	1	3	1.8	3	3	3	0.75	3	1.8	2.6	1	3	0.6
	CJ02	2	3	0	3	0.4	3	0	3	0.2	2.4	0.6	2.8	0
	CJ06	2	2.8	1.8	2.2	2.4	3	1	3	1	2.6	1.6	3	1.4
	CJ05	2	3	0.2	2.4	2.2	3	0	3	0.8	3	2.8	3	0.4
KJ int	KJ14	1	2.4	1.4	1.8	2.2	2.4	0	3	0.6	3	2.4	3	0.2
	KJ29	1	2.6	0.2	2.8	3	3	0	2.8	0.4	3	1.8	2.8	0.2
	KJ16	1	3	1.4	1.2	0.8	2.4	2.8	2	1.6	2.4	1.4	2.2	1.2
	KJ39	1	2	1.8	1.6	3	3	0.4	2.4	0.8	2	2.8	2.8	0.2
	KJ25	1	3	0	3	0.2	3	0.4	3	0.4	2.6	1	3	0.2
	KJ35	1	2.6	2.2	2.8	3	2.6	0.4	2.6	0.8	2.6	3	3	0.6
	KJ04	1	3	1.4	3	2.6	3	0	2.6	1.6	2.6	2.8	3	0.4
	KJ18	1	2.2	2	2	2.6	2	0.6	2.8	1	3	2.8	2.8	1.8
	KJ20	1	2	1.6	2.6	2.4	2.6	0.8	2.2	1.2	2.6	3	2.8	1.2
	KJ26	1	3	2.25	3	3	3	0	3	0	2.4	1.2	3	0
	KJ42	1	2.2	1.6	2	2	2.8	1	2	0.8	2.4	1.4	3	0.6
	KJ05	1	3	1	2.8	2.4	3	0.4	3	1.6	2.4	2.8	3	0.4
	KJ32	1	2	0	2.4	0	3	0	3	0	2.4	1.2	1.8	0
	KJ33	1	3	0	3	1.4	3	0	2.8	0.8	1.8	2.2	3	0
	KJ15	1	3	0.8	2.2	1.6	2.4	1	3	0	1.6	1	2.8	0
	KJ27	1	2.6	1.6	2	2.4	3	1.6	3	2	3	2.8	2.8	1.6
	KJ30	1	2.4	1.25	2.4	2	3	0.4	2.8	1	2.6	2.75	3	0.75
	KJ31	1	3	1.67	1.8	1.8	2.2	2.4	2.6	1	1.8	1.5	3	1
	KJ41	1	3	1.4	2.4	1.8	1.8	1.8	1.2	1.2	2.6	1.6	1.4	2
	KJ11	2	1.6	1.2	1.8	1.4	2	1.4	2	1.6	2	2.2	2.2	1.6
KJ adv	KJ23	1	2.6	0.8	2	1.2	3	0.8	2.4	1	2	2	2.8	0.4
	KJ36	1	2.6	0	3	0.4	3	0	3	0.4	2.8	2.2	3	0
	KJ06	1	3	1.2	3	0.6	3	0	3	0	2.4	1.8	3	0
	KJ13	2	2.8	1.4	3	1.4	3	1	2.6	0.8	2.8	2	3	1
	KJ02	2	2.8	0.8	2.6	2.2	3	0	3	1.4	2.8	2.8	3	0.8
	KJ12	2	2.6	2	3	1.6	3	1.2	3	0.4	2.8	1	3	0.6
	KJ22	1	2.4	1	1.8	1.6	2.2	1	2.4	1	2	0.4	2.4	0.2
	KJ08	2	3	0	2.8	1.8	3	0.2	3	1	2.8	3	3	0.4
	KJ09	2	1.8	0	2.4	1.4	3	0	2.8	0	2.8	2.4	2.6	0
	KJ38	1	2	0.6	2	1.8	3	0.4	3	0.6	3	3	3	0
	KJ01	2	2.4	1.8	2.8	3	3	0	3	0.6	2.8	3	3	0
	KJ10	2	3	0	3	2.6	3	0.6	3	0.6	2.6	2.4	3	0.2
	KJ34	1	2.8	0	2.6	2.4	2.6	0	2.8	1.2	2.6	2.2	2.8	0.2
	KJ03	2	2.6	0	2.4	2	3	0	2.8	2.2	2.6	2.2	3	1.4
	KJ07	1	2.6	0	3	2.2	3	0	3	0	2.4	1.6	2.6	1

	id	order	mean rating (scale 0 to 3)											
			1a	1b	1c	1d	2a	2b	3a	3b	3c	3d	4a	4b
EE	EE06	2	3	2.75	n/a	n/a	3	1.2	3	2	n/a	n/a	3	0.8
	EE12	2	2.8	3	n/a	n/a	3	1	3	2	n/a	n/a	3	1.2
	EE14	2	3	1.2	n/a	n/a	3	0.8	3	1.8	n/a	n/a	3	0.6
	EE15	2	2.4	1	n/a	n/a	3	0.2	3	1.8	n/a	n/a	3	0
	EE16	1	3	2.8	n/a	n/a	3	1.6	2.8	3	n/a	n/a	3	2
	EE17	1	3	3	n/a	n/a	3	0.8	3	3	n/a	n/a	3	1
	EE18	1	2.4	2.2	n/a	n/a	3	0	3	0	n/a	n/a	2.6	0.4
	EE19	1	3	1.2	n/a	n/a	3	0	3	0	n/a	n/a	3	0
	EE20	1	3	2.6	n/a	n/a	3	3	3	2.25	n/a	n/a	3	1.4
	EE21	1	3	1.2	n/a	n/a	3	0	3	0	n/a	n/a	3	0.8
	EE22	1	2.2	1.6	n/a	n/a	2.8	0.4	2.6	1.4	n/a	n/a	3	0.2
	EE23	1	3	0.4	n/a	n/a	3	0	2.8	0.2	n/a	n/a	3	0.2
	EE24	2	2.8	2.4	n/a	n/a	3	0.2	3	3	n/a	n/a	3	0
	EE25	2	3	2.4	n/a	n/a	3	1.6	3	3	n/a	n/a	3	2
	EE26	2	2.8	2.8	n/a	n/a	3	1.4	3	2.6	n/a	n/a	3	1.6
	EE27	2	1.8	2.6	n/a	n/a	3	2	3	1.2	n/a	n/a	2.8	0.6
	EE28	1	2.8	1.75	n/a	n/a	2.8	1	2.6	2.2	n/a	n/a	3	1
	EE29	2	2.2	1.6	n/a	n/a	3	0.4	2.8	1.2	n/a	n/a	3	0.4
	EE30	2	2.2	2.2	n/a	n/a	3	0.4	3	1.8	n/a	n/a	2.8	0.4
	EE31	2	2.2	0.8	n/a	n/a	2.8	0.8	2.6	1.8	n/a	n/a	2.8	0.8
	EE32	2	3	2	n/a	n/a	3	2	2.8	2	n/a	n/a	3	2
	EE33	2	2.6	1.8	n/a	n/a	3	1.6	3	1.6	n/a	n/a	3	0.6
	EE34	2	2.4	2	n/a	n/a	3	0	2.8	2.2	n/a	n/a	3	1
	EE35	2	2.8	1.8	n/a	n/a	3	1	3	1.8	n/a	n/a	2.8	1.4
CC	CC01	1	2.8	0.8	n/a	n/a	2.8	0.4	2.2	1.4	n/a	n/a	2.2	1
	CC15	1	2.8	0	n/a	n/a	2.4	0.4	3	0	n/a	n/a	2.6	0.2
	CC17	1	2.4	0.8	n/a	n/a	2.8	0.4	2.8	1.4	n/a	n/a	2.8	0.8
	CC18	1	2.2	0.6	n/a	n/a	3	0.6	2.8	0.8	n/a	n/a	2.2	1
	CC19	1	2.2	0	n/a	n/a	3	0	3	0	n/a	n/a	1.6	0.2
	CC20	1	2.2	0	n/a	n/a	3	0	2.4	0.6	n/a	n/a	3	0.2
	CC21	1	2.8	0.4	n/a	n/a	3	0	3	0	n/a	n/a	2	0.8
	CC22	1	2.2	0	n/a	n/a	2.8	0	3	0	n/a	n/a	2.6	0
	CC23	1	2.6	1	n/a	n/a	3	1.8	2.8	1.8	n/a	n/a	2.4	1.4
	CC24	1	2.4	0.4	n/a	n/a	2.6	0.4	2.75	0.4	n/a	n/a	2.2	0
	CC25	1	1.8	0.4	n/a	n/a	2	0	1.6	0	n/a	n/a	1.8	0
	CC26	1	2.4	0.2	n/a	n/a	3	0	2	0.6	n/a	n/a	2.6	0
	CC27	1	1.4	0	n/a	n/a	0	0	0	0	n/a	n/a	0.2	0
	CC28	1	2.4	0	n/a	n/a	3	0	2.6	0.2	n/a	n/a	2.6	0.6
	CC29	1	2.6	0.4	n/a	n/a	2.2	0	2.8	0.4	n/a	n/a	2.6	1.2
	CC31	1	1.6	0	n/a	n/a	2.8	0	1.2	0	n/a	n/a	1.6	0
	CC32	1	1.4	0	n/a	n/a	3	0.2	2.2	0.6	n/a	n/a	2	0
	CC33	1	2.8	0.4	n/a	n/a	3	0.4	2.6	0.8	n/a	n/a	2.6	0.4
	CC34	1	1.4	0	n/a	n/a	2.4	0.2	1.8	0	n/a	n/a	2.2	0.2
	CC35	1	2.8	0	n/a	n/a	2.2	0	2.6	0	n/a	n/a	2.6	0.6

Table 5F.1.iii: Mean ratings on test tokens by native Japanese, English and Chinese groups

index	JJ		CC		EE	
	mean	SD	mean	SD	mean	SD
1.1a	1.8	1.11	2.67	0.56	1.9	0.91
1.2a	2.35	0.93	2.83	0.38	2.5	0.69
1.3a	2.45	0.76	2.58	0.65	2.25	0.79
1.4a	2.3	0.86	2.88	0.34	2.4	0.75
1.5a	2.4	0.94	2.46	0.78	2.25	0.72
1.1b	0.75	0.97	2.5	0.59	0.35	0.59
1.2b	0.8	0.95	2.13	0.97	0.05	0.22
1.3b	0.75	0.85	2.09	1.08	0.25	0.55
1.4b	0.58	0.9	1.67	1.2	0.2	0.52
1.5b	0.55	1	1.42	1.14	0.5	0.69
1.1c	1.85	0.99	*n/a*	*n/a*	*n/a*	*n/a*
1.2c	2	0.97	*n/a*	*n/a*	*n/a*	*n/a*
1.3c	2.15	0.88	*n/a*	*n/a*	*n/a*	*n/a*
1.4c	2.05	0.69	*n/a*	*n/a*	*n/a*	*n/a*
1.5c	2	0.97	*n/a*	*n/a*	*n/a*	*n/a*
1.1d	2.3	0.86	*n/a*	*n/a*	*n/a*	*n/a*
1.2d	2.05	1.03	*n/a*	*n/a*	*n/a*	*n/a*
1.3d	2.6	0.68	*n/a*	*n/a*	*n/a*	*n/a*
1.4d	2.55	0.83	*n/a*	*n/a*	*n/a*	*n/a*
1.5d	1.9	0.97	*n/a*	*n/a*	*n/a*	*n/a*
2.1a	2.45	0.6	2.96	0.2	2.45	0.89
2.2a	2.6	0.68	3	0	2.75	0.72
2.3a	2.45	0.89	2.96	0.2	2.6	0.94
2.4a	2.45	0.83	3	0	2.8	0.7
2.5a	2.75	0.64	2.96	0.2	2.4	0.94
2.1b	0.65	0.81	0.79	0.78	0.2	0.52
2.2b	0.55	0.76	0.71	0.91	0.3	0.57
2.3b	0.65	0.75	0.79	0.83	0.15	0.49
2.4b	0.5	0.76	0.96	1.04	0.3	0.47
2.5b	0.65	0.75	1.21	1.02	0.25	0.55
3.1a	2.75	0.72	2.96	0.2	1.95	1.13
3.2a	2.9	0.45	3	0	2.45	1
3.3a	2.9	0.31	2.83	0.38	2.35	0.88
3.4a	3	0	2.96	0.2	2.6	0.82
3.5a	2.75	0.72	2.79	0.41	2.4	0.n/a
3.1b	0.7	0.8	1.58	1.02	0.25	0.55
3.2b	0.9	1.07	1.7	1.06	0.55	0.89
3.3b	0.55	0.76	1.71	1.16	0.4	0.68
3.4b	0.65	0.81	1.96	1.08	0.65	1.04
3.5b	0.7	0.92	1.75	0.68	0.4	0.68
3.1c	2.95	0.22	*n/a*	*n/a*	*n/a*	*n/a*
3.2c	2.89	0.46	*n/a*	*n/a*	*n/a*	*n/a*
3.3c	2.55	0.6	*n/a*	*n/a*	*n/a*	*n/a*
3.4c	2.15	1.04	*n/a*	*n/a*	*n/a*	*n/a*
3.5c	2.53	0.61	*n/a*	*n/a*	*n/a*	*n/a*
3.1d	2.4	0.82	*n/a*	*n/a*	*n/a*	*n/a*
3.2d	2.2	1.06	*n/a*	*n/a*	*n/a*	*n/a*
3.3d	2.4	0.75	*n/a*	*n/a*	*n/a*	*n/a*
3.4d	2.85	0.67	*n/a*	*n/a*	*n/a*	*n/a*
3.5d	2.47	0.84	*n/a*	*n/a*	*n/a*	*n/a*
4.1a	2.25	0.85	2.92	0.41	1.65	0.75
4.2a	2.95	0.22	3	0	2.5	0.83

index	JJ		CC		EE	
	mean	SD	mean	SD	mean	SD
4.3a	2.85	0.49	2.83	0.38	2.75	0.72
4.4a	2.75	0.72	3	0	2.53	0.7
4.5a	2.7	0.57	3	0	1.7	0.92
4.1b	0.5	0.89	0.78	0.9	0.3	0.57
4.2b	0.95	0.89	1.08	0.97	0.65	0.93
4.3b	0.65	0.93	0.67	0.82	0.5	0.69
4.4b	1	0.86	0.83	0.78	0.35	0.67
4.5b	0.55	0.94	0.79	0.78	0.35	0.59
T1D01	3	0	2.96	0.2	2.43	0.9
T1D02	1.6	1.27	0.08	0.28	0.7	1.02
T1D03	2.52	0.81	2.83	0.38	2.75	0.44
T1D04	2.86	0.36	2.92	0.28	1.5	1.18
T1D05	2.71	0.46	2.88	0.34	2.75	0.53
T1D06	0.55	0.89	1.21	1.02	1.21	1.14
T1D07	0.14	0.48	0.29	0.55	0.58	0.83
T1D08	2.86	0.48	2.92	0.41	2.88	0.45
T1D09	0.1	0.3	0	0	0.21	0.51
T1D10	0.05	0.22	0	0	0.21	0.51
T1D11	0.1	0.3	0.04	0.2	0.26	0.54
T1D12	2.14	0.85	3	0	1.74	1.05
T1D13	2.57	0.81	2.88	0.45	2.46	0.98
T1D14	0.57	1.08	0.08	0.28	0.29	0.69
T2D01	2.95	0.22	3	0	2.74	0.45
T2D02	0.81	1.08	0.04	0.2	0.33	0.7
T2D03	2.33	0.86	3	0	2.67	0.7
T2D04	2.86	0.36	2.92	0.28	1.88	1.12
T2D05	2.57	0.81	2.79	0.51	2.54	0.72
T2D06	0.35	0.59	0.67	0.87	0.92	0.97
T2D07	0.24	0.44	0.29	0.69	2.61	0.66
T2D08	2.95	0.22	3	0	2.91	0.29
T2D09	0.14	0.48	0.04	0.2	0.79	0.98
T2D10	0.05	0.22	0.13	0.61	0.13	0.34
T2D11	0.1	0.3	0.04	0.2	0.55	0.74
T2D12	2.43	0.68	3	0	2.09	0.85
T2D13	2.62	0.5	2.88	0.45	2.65	0.57
T2D14	0.19	0.68	3	0	2.25	0.44

Table 5F.1.iv: Mean ratings on test tokens by L2 groups

index	EJ int		EJ adv		CJ int		CJ adv		KJ int		KJ adv	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
1.1a	2.65	0.61	2.89	0.33	2.67	0.52	2.78	0.67	2.47	0.9	2.73	0.46
1.2a	2.83	0.38	2.78	0.67	2.5	0.55	3	0	2.75	0.55	2.8	0.56
1.3a	2.83	0.38	2.78	0.67	1.67	1.37	3	0	2.45	0.83	2.67	0.62
1.4a	2.78	0.73	2.89	0.33	1.5	1.38	2.78	0.44	2.75	0.72	2.6	0.74
1.5a	2.61	0.61	2.78	0.44	2	1.55	2.78	0.44	2.45	0.69	2.2	0.77
1.1b	1.67	0.84	1.11	1.27	1.33	1.51	0.89	0.93	1.2	1.15	0.73	1.03
1.2b	1.33	0.97	1.44	1.13	0.8	1.3	1	1	1.2	1.24	0.33	0.82
1.3b	2	0.79	1.44	1.24	0.67	1.21	1.11	1.17	1.89	1.29	0.6	0.99
1.4b	1.72	0.83	1.56	1.42	1.5	1.05	0.67	0.87	0.89	1.02	1.13	1.41
1.5b	1.39	1.2	1.33	1.41	1.67	1.21	1.67	1.22	0.89	0.94	0.4	0.74
1.1c	2.56	0.62	2.67	0.71	2.67	0.82	2.33	0.71	2.55	0.89	2.73	0.46
1.2c	2.78	0.43	2.67	0.71	2.17	0.98	3	0	2.25	0.91	2.67	0.82
1.3c	2.65	0.79	2.56	0.88	1.5	1.38	2.44	0.88	2.55	0.69	2.47	0.83
1.4c	2.67	0.49	2.89	0.33	2.2	0.84	2.89	0.33	2.25	0.97	2.8	0.41
1.5c	2.61	0.78	2.67	0.71	1.67	1.21	2.56	0.73	2.05	0.94	2.47	0.74
1.1d	2.56	0.62	1.67	1.12	1.67	1.21	1.56	1.33	2.45	0.89	2	1.07
1.2d	2.33	0.77	1.89	1.17	1.8	1.3	1.89	1.05	1.75	1.29	0.93	0.96
1.3d	2.61	0.61	1.67	1.22	1.5	1.38	2.67	0.5	2.05	1.28	2.33	0.9
1.4d	2.67	0.59	1.44	1.13	2	1.26	2.33	1.12	2	1.12	1.8	1.15
1.5d	2.22	0.88	1.56	1.33	2.17	1.33	2.44	1.01	1.58	1.22	1.67	0.9
2.1a	2.94	0.24	3	0	2.17	1.17	3	0	2.5	0.89	2.87	0.35
2.2a	3	0	3	0	2.33	0.82	3	0	2.9	0.31	2.93	0.26
2.3a	2.72	0.46	3	0	2.33	1.21	2.67	1	2.45	0.94	2.87	0.35
2.4a	2.89	0.47	2.89	0.33	2	1.26	2.67	1	2.6	0.94	3	0
2.5a	2.83	0.51	3	0	1.5	1.64	2.78	0.67	2.85	0.37	2.93	0.26
2.1b	1.78	1	1.11	1.17	0.5	0.55	0.89	0.93	0.55	1	0.27	0.46
2.2b	1.61	0.98	1.11	1.17	0.5	0.84	0.38	0.74	0.8	1.15	0.53	0.83
2.3b	1.67	0.77	1.11	1.17	0.83	1.17	0.13	0.35	0.9	1.02	0.2	0.41
2.4b	2.06	0.73	1.22	1.09	0.5	1.22	0.67	0.71	0.8	1.01	0.2	0.41
2.5b	2	0.84	1.33	1.12	0.5	0.55	0.56	1.01	0.8	1.11	0.53	0.83
3.1a	2.94	0.24	3	0	1.83	1.17	3	0	2.55	0.69	2.87	0.35
3.2a	2.89	0.32	2.78	0.67	2.5	0.84	2.78	0.67	2.45	1.05	2.93	0.26
3.3a	2.72	0.46	2.89	0.33	2.33	0.52	2.89	0.33	2.55	0.83	2.73	0.46
3.4a	2.72	0.75	3	0	2	1.1	3	0	2.8	0.7	2.87	0.35
3.5a	2.78	0.43	3	0	2.17	1.17	2.78	0.44	2.6	0.68	2.87	0.35
3.1b	1.83	0.92	1	0.87	0.83	1.33	1.44	1.33	0.85	0.88	0.8	0.68
3.2b	2	0.91	1.44	1.24	0.67	0.82	0.67	0.5	1.05	1.05	0.8	0.86
3.3b	1.94	0.73	1.56	1.24	0.67	0.82	0.22	0.44	0.6	0.82	0.53	0.64
3.4b	2	1.19	1.33	1.22	1.5	1.38	0.78	1.09	1.1	1.17	0.73	0.96
3.5b	1.83	1.04	1.22	0.97	0.83	1.33	1	1	1	1.12	0.87	0.92
3.1c	2.88	0.33	3	0	2.33	0.82	3	0	2.63	0.76	3	0
3.2c	2.72	0.75	3	0	2	1.26	2.78	0.67	2.95	0.22	2.93	0.26
3.3c	2.56	0.62	2.67	0.71	2.5	0.84	2.89	0.33	2.2	1.15	2.2	0.77
3.4c	1.78	0.94	2.33	1.12	2.33	0.82	1.89	1.05	2.55	1	2.27	1.03
3.5c	2.29	1.16	2.78	0.67	2.17	1.17	2.89	0.33	1.85	1.27	2.67	0.62
3.1d	1.88	0.86	1.78	1.09	2.5	0.55	1.67	1.22	2	1.3	2.07	1.03
3.2d	1.94	0.73	1.78	1.09	2.33	0.82	1.56	0.88	1.85	1.35	1.87	1.13
3.3d	1.5	0.99	1.44	1.13	2.33	0.82	1.67	0.87	2.32	0.89	2.13	0.92
3.4d	2.72	0.46	2.89	0.33	2	1.26	2.33	1.12	2.45	1.1	2.4	1.24
3.5d	1.72	0.83	1.89	1.05	2.33	0.52	2.22	1.09	1.79	1.4	2.2	0.94
4.1a	2.78	0.55	2.89	0.33	2	1.26	3	0	2.45	0.76	2.87	0.35
4.2a	2.83	0.71	3	0	1.83	1.47	2.78	0.44	2.7	0.8	3	0
4.3a	2.76	0.44	2.89	0.33	2.5	0.84	2.89	0.33	2.8	0.7	2.8	0.56

index	EJ int		EJ adv		CJ int		CJ adv		KJ int		KJ adv	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
4.4a	2.75	0.58	3	0	1.83	1.47	3	0	2.85	0.67	3	0
4.5a	2.94	0.24	3	0	1.83	1.17	2.78	0.44	2.8	0.7	2.73	0.59
4.1b	1.44	1.1	1.22	0.97	0.67	1.21	0.78	0.97	0.75	1.16	0.47	0.92
4.2b	1.33	0.91	1.11	0.93	0.33	0.82	1.11	1.27	0.85	1.04	0.47	0.64
4.3b	1.33	0.97	1	0.87	0.83	0.98	0.89	0.93	0.7	0.98	0.33	0.62
4.4b	1.56	0.98	1	0.87	0.5	0.55	0.78	0.97	0.45	0.83	0.47	0.74
4.5b	0.94	1	0.89	0.93	0.83	1.33	0.56	0.53	0.74	0.99	0.33	0.62
T1D01	2.85	0.37	3	0	2.83	0.41	3	0	2.91	0.29	2.87	0.35
T1D02	2.05	0.83	2.22	1.2	1.57	1.51	2.1	0.99	1.68	1.29	2.13	1.06
T1D03	2.75	0.44	2.89	0.33	2.43	0.79	2.7	0.67	2.55	0.86	2.53	0.64
T1D04	2.4	0.6	2.67	1	1.71	1.11	2.1	0.99	2.62	0.67	2.5	0.65
T1D05	2.22	1.06	2.56	0.73	2.57	0.53	2.9	0.32	2.32	0.78	2.67	0.62
T1D06	0.75	1.12	0.56	1.01	0.86	1.21	1.4	0.97	1	1.15	0.6	0.83
T1D07	0.41	1	0.44	0.53	0.29	0.76	0.6	0.97	0.18	0.66	0.13	0.35
T1D08	2.95	0.22	3	0	2.71	0.49	2.9	0.32	2.95	0.21	3	0
T1D09	0.1	0.31	0	0	0.43	1.13	0.3	0.67	0.1	0.3	0	0
T1D10	0.2	0.7	0	0	0.43	1.13	0.4	0.97	0.05	0.21	0	0
T1D11	0.25	0.79	0.11	0.33	0.14	0.38	0.3	0.95	0.14	0.47	0.13	0.35
T1D12	2.75	0.55	3	0	2.71	0.49	2.8	0.42	2.43	0.68	2.67	0.49
T1D13	2.68	0.75	2.75	0.71	2.71	0.49	2.1	1.2	2.82	0.5	2.53	0.92
T1D14	0.35	0.75	0.22	0.67	0.29	0.76	0.4	0.97	0.23	0.69	0.13	0.52
T2D01	2.5	0.83	2.67	0.71	2	1.15	2.2	1.32	2.73	0.77	2.93	0.26
T2D02	2.9	0.31	1.56	1.42	1.6	1.52	1.7	1.16	1.32	1.25	0.87	0.99
T2D03	2.75	0.44	2.78	0.44	2.29	0.76	2.7	0.48	2.73	0.55	2.6	0.51
T2D04	2.95	0.22	3	0	2.43	0.79	3	0	2.77	0.53	2.8	0.41
T2D05	2.6	0.5	2.75	0.71	2.43	1.13	2	1.22	2.18	0.85	2.47	0.83
T2D06	1.3	1.13	0.67	1	1.14	1.35	0.3	0.48	0.91	1.02	0.4	0.63
T2D07	1.11	0.99	0.89	0.78	0	0	1	1.25	0.67	1.02	0.2	0.41
T2D08	2.9	0.45	3	0	2.86	0.38	2.8	0.42	2.95	0.21	2.93	0.26
T2D09	0.05	0.22	0	0	0.14	0.38	0.6	1.07	0.09	0.29	0.07	0.26
T2D10	0.25	0.72	0	0	0.29	0.76	0.4	0.97	0.05	0.21	0	0
T2D11	0.35	0.81	0.11	0.33	0.29	0.49	0.3	0.67	0.32	0.65	0.2	0.56
T2D12	2.89	0.32	3	0	2.71	0.49	2.5	1.08	2.41	0.91	2.73	0.46
T2D13	2.7	0.73	2.89	0.33	1.14	1.46	2	1.41	2.82	0.39	2.4	1.12
T2D14	0.25	0.79	0.13	0.35	0.14	0.38	0.3	0.95	0.05	0.21	0.07	0.26

Table 5F.2.i: Mean ratings on Korean test types by participant[221]

id	order	mean rating (scale 0 to 3)							
		K1a	K1b	K1c	K1d	K2a	K2b	K2c	K2d
KK01	1	1.6	0.6	2	1.6	1.8	1.2	2.4	1.2
KK02	1	2.4	0.6	3	1.2	2.8	0.6	2.4	0.6
KK03	1	3	0.4	2.6	1.4	3	1.2	3	2.4
KK04	1	2.8	1.4	3	2.4	3	1.2	3	1.8
KK06	1	1.6	0.6	1.8	1.2	2.2	0.4	1.8	0.6
KK08	1	1	0.2	1.2	0.6	1.2	1	0	0
KK09	1	0.8	1	1.4	1	1.8	1.8	0.6	0.6
KK10	1	2.5	0.8	2.8	2	3	0.6	3	1.2
KK11	1	1	0.6	1.4	0.8	1.4	1.2	0	1.2
KK12	1	1.2	0	2.2	0.8	1.6	0.8	1.8	1.8
KK13	1	2.6	0.2	3	2.8	3	0.6	2.4	3
KK14	1	2.8	0	3	0.6	2.6	0.6	2.4	0.6
KK15	1	2.6	0.4	2.8	0.4	3	0.4	2.4	0.6
KK16	1	2.2	0.2	2.2	1	2.2	1.2	2.4	1.2
KK17	1	1.8	0.2	2	2.6	2.2	1.2	0.6	3
KK18	1	2.8	0	2.4	0.2	2.6	0.4	2.4	1.8
KK19	1	1.8	0	3	0.6	2.4	1.8	1.8	1.2
KK20	1	2	0.6	2.4	1	2.2	1.4	1.2	1.8
KK21	1	2	1.2	2	2	2.2	2	3	3
KK22	1	1.4	1.4	1.8	1.8	2.8	1.8	3	2.4
KK24	1	2.4	0.4	1	0.4	2	0.2	2.4	1.2
KK25	1	2.8	0	3	0.8	2.6	0	2.4	0.6

[221] See Appendix 3 for proficiency scores for learners and for age and gender of all participants.

Appendix 5F.3: Repeated measures ANOVA for non-scrambled QP-QP sentences[222]

Table 5F.3.i: Descriptive statistics

Group	Type 1a	Type 1b	Type 2a	Type 2b	Type 3a	Type 3b	Type 4a	Type 4b
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
JJ (n=20)	2.26 (0.70)	0.69 (0.72)	2.54 (0.54)	0.60 (0.54)	2.86 (0.27)	0.70 (0.62)	2.70 (0.31)	0.73 (0.66)
EJ int (n=18)	2.74 (0.28)	1.61 (0.63)	2.88 (0.16)	1.76 (0.67)	2.81 (0.27)	1.92 (0.77)	2.80 (0.29)	1.32 (0.71)
EJ adv (n=9)	2.82 (0.39)	1.38 (1.12)	2.98 (0.07)	1.16 (0.88)	2.93 (0.14)	1.31 (1.02)	2.96 (0.09)	1.04 (0.82)
CJ int (n=6)	2.07 (0.98)	1.19 (0.77)	2.07 (1.04)	0.60 (0.55)	2.17 (0.85)	0.90 (0.60)	2.00 (0.98)	0.63 (0.61)
CJ adv (n=9)	2.87 (0.17)	1.07 (0.78)	2.82 (0.37)	0.71 (0.44)	2.89 (0.20)	0.82 (0.53)	2.89 (0.20)	0.82 (0.60)
KJ int (n=20)	2.58 (0.45)	1.24 (0.70)	2.66 (0.41)	0.79 (0.70)	2.59 (0.49)	0.92 (0.58)	2.72 (0.46)	0.70 (0.66)
KJ adv (n=15)	2.60 (0.35)	0.64 (0.71)	2.92 (0.22)	0.51 (0.39)	2.85 (0.22)	0.75 (0.59)	2.88 (0.20)	0.41 (0.45)
EE (n=24)	2.68 (0.36)	1.96 (0.73)	2.98 (0.68)	0.93 (0.66)	2.91 (0.14)	1.74 (0.94)	2.95 (0.11)	0.85 (0.63)
CC (n=20)	2.26 (0.49)	0.27 (0.33)	2.60 (0.69(	0.41 (0.33)	2.36 (0.75)	0.45 (0.55)	2.22 (0.61)	0.43 (0.46)

Table 5F.3.ii: Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2995.394	1	2995.394	4033.722	.000
Group	92.740	8	11.593	15.611	.000
Error	98.022	132	.743		

[222] Sphericity is assumed, as Mauchly's test of sphericity was non-significant. Equality of variance not assumed, due to significant results in Levene's test:

Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
Type 1A	4.529	8	132	.000
Type 1B	2.792	8	132	.007
Type 2A	11.348	8	132	.000
Type 2B	2.195	8	132	.032
Type 3A	6.941	8	132	.000
Type 3B	1.590	8	132	.134
Type 4A	9.582	8	132	.000
Type 4B	.602	8	132	.775

Table 5F.3.iii: Tests of Within Subjects Contrasts

Measure: MEASURE_1

Source	Subj QP	Obj QP	Scope	Type III Sum of Squares	df	Mean Square	F	Sig.
Subj QP	Linear			.002	1	.002	.013	.911
Subj QP * Group	Linear			2.525	8	.316	1.985	.053
Error(Subj QP)	Linear			20.983	132	.159		
Obj QP		Linear		2.652	1	2.652	23.095	.000
Obj QP * Group		Linear		6.534	8	.817	7.114	.000
Error(Obj QP)		Linear		15.155	132	.115		
Scope			Linear	679.923	1	679.923	827.148	.000
Scope * Group			Linear	29.930	8	3.741	4.551	.000
Error(Scope)			Linear	108.505	132	.822		
Subj QP * Obj QP	Linear	Linear		.577	1	.577	3.936	.049
Subj QP * Obj QP * Group	Linear	Linear		2.778	8	.347	2.367	.021
Error(Subj QP*Obj QP)	Linear	Linear		19.364	132	.147		
Subj QP * Scope	Linear		Linear	.872	1	.872	4.392	.038
Subj QP * Scope * Group	Linear		Linear	2.057	8	.257	1.295	.251
Error(Subj QP*Scope)	Linear		Linear	26.203	132	.199		
Obj QP * Scope		Linear	Linear	7.427	1	7.427	50.772	.000
Obj QP * Scope * Group		Linear	Linear	9.034	8	1.129	7.719	.000
Error(Obj QP*Scope)		Linear	Linear	19.310	132	.146		
Subj QP * Obj QP * Scope	Linear	Linear	Linear	.598	1	.598	3.662	.058
Subj QP * Obj QP * Scope * Group	Linear	Linear	Linear	1.983	8	.248	1.518	.157
Error(Subj QP*Obj QP * Scope)	Linear	Linear	Linear	21.544	132	.163		

Table 5F.3.iv:
Between-groups post hoc multiple comparisons (Games-Howell), Type 1b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ int	-.9239(*)	.21862	.004	-1.6447	-.2031
	EJ adv	-.6878	.40478	.737	-2.2179	.8424
	CJ int	-.5017	.35446	.864	-1.9601	.9567
	CJ adv	-.3767	.30553	.935	-1.4812	.7278
	KJ int	-.5483	.22461	.292	-1.2865	.1899
	KJ adv	.0500	.24404	1.000	-.7638	.8638
	EE	-1.2725(*)	.21881	.000	-1.9887	-.5563
	CC	.4200	.17577	.329	-.1722	1.0122
EJ int	EJ adv	.2361	.40061	.999	-1.2903	1.7625
	CJ int	.4222	.34970	.932	-1.0382	1.8827
	CJ adv	.5472	.29999	.668	-.5483	1.6428
	KJ int	.3756	.21702	.724	-.3400	1.0911
	KJ adv	.9739(*)	.23707	.008	.1793	1.7685
	EE	-.3486	.21100	.770	-1.0409	.3437
	CC	1.3439(*)	.16595	.000	.7816	1.9062
EJ adv	CJ int	.1861	.48820	1.000	-1.6078	1.9801
	CJ adv	.3111	.45393	.998	-1.3297	1.9520
	KJ int	.1394	.40392	1.000	-1.3898	1.6687
	KJ adv	.7378	.41503	.695	-.8074	2.2830
	EE	-.5847	.40072	.852	-2.1102	.9408
	CC	1.1078	.37894	.204	-.4074	2.6229
CJ int	CJ adv	.1250	.40969	1.000	-1.4277	1.6777
	KJ int	-.0467	.35348	1.000	-1.5052	1.4119
	KJ adv	.5517	.36613	.828	-.9125	2.0158
	EE	-.7708	.34982	.482	-2.2296	.6879
	CC	.9217	.32464	.275	-.5736	2.4170
CJ adv	KJ int	-.1717	.30439	1.000	-1.2741	.9307
	KJ adv	.4267	.31900	.905	-.7104	1.5638
	EE	-.8958	.30013	.150	-1.9894	.1977
	CC	.7967	.27037	.191	-.2651	1.8584
KJ int	KJ adv	.5983	.24261	.287	-.2112	1.4078
	EE	-.7242(*)	.21721	.043	-1.4350	-.0134
	CC	.9683(*)	.17378	.000	.3832	1.5534
KJ adv	EE	-1.3225(*)	.23724	.000	-2.1135	-.5315
	CC	.3700	.19825	.642	-.3228	1.0628
EE	CC	1.6925(*)	.16620	.000	1.1415	2.2435

Based on observed means.
* The mean difference is significant at the .05 level.

Table 5F.3.v:
Between-groups post hoc multiple comparisons (Games-Howell), Type 2b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ int	-1.1556(*)	.19758	.000	-1.8111	-.5000
	EJ adv	-.5556	.31739	.710	-1.7626	.6515
	CJ int	.0000	.25492	1.000	-1.0376	1.0376
	CJ adv	-.1111	.18855	.999	-.7683	.5460
	KJ int	-.1900	.19654	.987	-.8384	.4584
	KJ adv	.0933	.15669	1.000	-.4262	.6128
	EE	-.3333	.18116	.656	-.9255	.2588
	CC	.1900	.14103	.909	-.2787	.6587
EJ int	EJ adv	.6000	.33338	.682	-.6281	1.8281
	CJ int	1.1556(*)	.27457	.029	.1031	2.2080
	CJ adv	1.0444(*)	.21437	.002	.3125	1.7764
	KJ int	.9656(*)	.22143	.003	.2353	1.6958
	KJ adv	1.2489(*)	.18696	.000	.6219	1.8759
	EE	.8222(*)	.20790	.009	.1376	1.5068
	CC	1.3456(*)	.17405	.000	.7549	1.9362
EJ adv	CJ int	.5556	.37025	.837	-.8040	1.9151
	CJ adv	.4444	.32811	.895	-.7827	1.6716
	KJ int	.3656	.33276	.964	-.8608	1.5919
	KJ adv	.6489	.31089	.527	-.5534	1.8512
	EE	.2222	.32392	.998	-.9912	1.4357
	CC	.7456	.30330	.355	-.4527	1.9439
CJ int	CJ adv	-.1111	.26814	1.000	-1.1698	.9476
	KJ int	-.1900	.27382	.998	-1.2400	.8600
	KJ adv	.0933	.24678	1.000	-.9477	1.1343
	EE	-.3333	.26300	.919	-1.3722	.7055
	CC	.1900	.23715	.992	-.8608	1.2408
CJ adv	KJ int	-.0789	.21341	1.000	-.8051	.6473
	KJ adv	.2044	.17740	.956	-.4294	.8382
	EE	-.2222	.19934	.966	-.9049	.4605
	CC	.3011	.16373	.660	-.3050	.9072
KJ int	KJ adv	.2833	.18586	.836	-.3356	.9023
	EE	-.1433	.20691	.999	-.8215	.5349
	CC	.3800	.17287	.434	-.2012	.9612
KJ adv	EE	-.4267	.16952	.257	-.9847	.1314
	CC	.0967	.12573	.997	-.3259	.5192
EE	CC	.5233(*)	.15516	.042	.0109	1.0357

Based on observed means.
* The mean difference is significant at the .05 level.

Table 5F.3.vi:
Between-groups post hoc multiple comparisons (Games-Howell), Type 3b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ int	-1.2222(*)	.22902	.000	-1.9819	-.4626
	EJ adv	-.6111	.36763	.756	-2.0080	.7858
	CJ int	-.2000	.28309	.997	-1.3362	.9362
	CJ adv	-.1222	.22601	1.000	-.9140	.6696
	KJ int	-.2200	.19079	.961	-.8472	.4072
	KJ adv	-.0467	.20710	1.000	-.7361	.6427
	EE	-1.0438(*)	.23666	.002	-1.8191	-.2684
	CC	.2500	.18659	.912	-.3637	.8637
EJ int	EJ adv	.6111	.38556	.797	-.8096	2.0318
	CJ int	1.0222	.30601	.099	-.1369	2.1814
	CJ adv	1.1000(*)	.25413	.007	.2294	1.9706
	KJ int	1.0022(*)	.22339	.003	.2592	1.7452
	KJ adv	1.1756(*)	.23747	.001	.3846	1.9666
	EE	.1785	.26365	.999	-.6860	1.0429
	CC	1.4722(*)	.21982	.000	.7396	2.2049
EJ adv	CJ int	.4111	.41994	.982	-1.1324	1.9546
	CJ adv	.4889	.38378	.921	-.9385	1.9163
	KJ int	.3911	.36415	.966	-1.0026	1.7849
	KJ adv	.5644	.37295	.829	-.8398	1.9687
	EE	-.4326	.39014	.963	-1.8581	.9928
	CC	.8611	.36197	.381	-.5309	2.2531
CJ int	CJ adv	.0778	.30376	1.000	-1.0980	1.2535
	KJ int	-.0200	.27856	1.000	-1.1554	1.1154
	KJ adv	.1533	.28997	1.000	-.9902	1.2968
	EE	-.8438	.31177	.241	-2.0063	.3188
	CC	.4500	.27570	.768	-.6856	1.5856
CJ adv	KJ int	-.0978	.22030	1.000	-.8762	.6807
	KJ adv	.0756	.23457	1.000	-.7439	.8950
	EE	-.9215(*)	.26103	.036	-1.8044	-.0387
	CC	.3722	.21667	.729	-.3982	1.1426
KJ int	KJ adv	.1733	.20085	.993	-.4970	.8437
	EE	-.8238(*)	.23121	.025	-1.5824	-.0651
	CC	.4700	.17963	.214	-.1205	1.0605
KJ adv	EE	-.9971(*)	.24485	.007	-1.8031	-.1911
	CC	.2967	.19687	.843	-.3618	.9551
EE	CC	1.2938(*)	.22776	.000	.5454	2.0421

Based on observed means.
* The mean difference is significant at the .05 level.

Table 5F.3.vii:
Between-groups post hoc multiple comparisons (Games-Howell), Type 4b

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ int	-.5922	.22247	.199	-1.3271	.1427
	EJ adv	-.3144	.31153	.978	-1.4606	.8318
	CJ int	.0967	.29015	1.000	-1.0573	1.2506
	CJ adv	-.0922	.24947	1.000	-.9736	.7892
	KJ int	.0325	.20859	1.000	-.6531	.7181
	KJ adv	.3167	.18776	.750	-.3061	.9394
	EE	-.1200	.19570	.999	-.7615	.5215
	CC	.3000	.17980	.760	-.2950	.8950
EJ int	EJ adv	.2778	.32103	.992	-.8857	1.4413
	CJ int	.6889	.30032	.423	-.4743	1.8521
	CJ adv	.5000	.26123	.613	-.4117	1.4117
	KJ int	.6247	.22253	.149	-.1104	1.3598
	KJ adv	.9089(*)	.20313	.003	.2296	1.5882
	EE	.4722	.21049	.402	-.2239	1.1683
	CC	.8922(*)	.19579	.002	.2366	1.5478
EJ adv	CJ int	.4111	.37115	.962	-.9561	1.7783
	CJ adv	.2222	.34030	.999	-1.0034	1.4479
	KJ int	.3469	.31157	.961	-.7993	1.4932
	KJ adv	.6311	.29803	.507	-.4985	1.7607
	EE	.1944	.30310	.999	-.9392	1.3280
	CC	.6144	.29307	.521	-.5095	1.7384
CJ int	CJ adv	-.1889	.32084	.999	-1.4087	1.0310
	KJ int	-.0642	.29020	1.000	-1.2182	1.0898
	KJ adv	.2200	.27560	.993	-.9349	1.3749
	EE	-.2167	.28107	.995	-1.3676	.9342
	CC	.2033	.27024	.995	-.9540	1.3606
CJ adv	KJ int	.1247	.24953	1.000	-.7568	1.0062
	KJ adv	.4089	.23239	.705	-.4403	1.2580
	EE	-.0278	.23886	1.000	-.8852	.8297
	CC	.3922	.22600	.718	-.4446	1.2291
KJ int	KJ adv	.2842	.18783	.842	-.3389	.9072
	EE	-.1525	.19577	.997	-.7942	.4892
	CC	.2675	.17987	.854	-.3277	.8627
KJ adv	EE	-.4367	.17340	.257	-1.0082	.1348
	CC	-.0167	.15523	1.000	-.5339	.5006
EE	CC	.4200	.16475	.239	-.1189	.9589

Based on observed means.
* The mean difference is significant at the .05 level.

5F.4: Repeated measures ANOVA for scrambled QP-QP sentences[223]

Table 5F.4.i: Descriptive statistics

Group	Type 1a[a] (S>O)	Type 1b[a] (O>S)	Type 1c[a] (S>O)	Type 1d[a] (O>S)	Type 3a[b] (S>O)	Type 3b[b] (O>S)	Type 3c[b] (S>O)	Type 3d[b] (O>S)
	(canonical: SOV)		(scrambled: OSV)		(canonical: SOV)		(scrambled: OSV)	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
JJ (n=20)	2.26 (0.70)	0.69 (0.72)	2.01 (0.71)	2.29 (0.56)	2.86 (0.27)	0.70 (0.62)	2.59 (0.35)	2.43 (0.56)
EJ int (n=18)	2.74 (0.28)	1.61 (0.63)	2.65 (0.35)	2.48 (0.54)	2.81 (0.27)	1.92 (0.77)	2.45 (0.41)	1.96 (0.52)
EJ adv (n=9)	2.82 (0.39)	1.38 (1.12)	2.69 (0.65)	1.64 (1.16)	2.93 (0.14)	1.31 (1.02)	2.76 (0.30)	1.96 (0.85)
CJ int (n=6)	2.07 (0.98)	1.19 (0.77)	2.03 (0.91)	1.83 (1.08)	2.17 (0.85)	0.90 (0.60)	2.27 (0.79)	2.30 (0.43)
CJ adv (n=9)	2.87 (0.17)	1.07 (0.78)	2.64 (0.40)	2.18 (0.72)	2.89 (0.20)	0.82 (0.53)	2.69 (0.23)	1.89 (0.72)
KJ int (n=20)	2.58 (0.45)	1.24 (0.70)	2.33 (0.54)	1.98 (0.88)	2.59 (0.49)	0.92 (0.58)	2.44 (0.42)	2.08 (0.74)
KJ adv (n=15)	2.60 (0.35)	0.64 (0.71)	2.63 (0.42)	1.75 (0.70)	2.85 (0.22)	0.75 (0.59)	2.61 (0.30)	2.13 (0.74)

Table 5F.4.ii: Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2745.211	1	2745.211	3664.313	.000
Group	15.563	6	2.594	3.462	.004
Error	67.426	90	.749		

[223] Sphericity is assumed, as Mauchly's test of sphericity was non-significant. Equality of variance not assumed, due to significant results in Levene's test:

Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
Type 1A	5.144	6	90	.000
Type 1B	1.544	6	90	.173
Type 1C	4.437	6	90	.001
Type 1D	2.449	6	90	.031
Type 3A	5.813	6	90	.000
Type 3B	1.754	6	90	.118
Type 3C	3.418	6	90	.004
Type 3D	1.452	6	90	.204

Table 5F.4.iii: Tests of Within Subjects Contrasts

Measure: MEASURE_1

Source*	Subj QP	Scram-bling	Scope	Type III Sum of Squares	df	Mean Square	F	Sig.
Subj QP	Linear			.887	1	.887	3.118	.081
Subj QP * Group	Linear			5.004	6	.834	2.934	.012
Error(Subj QP)	Linear			25.587	90	.284		
Scrambling		Linear		26.990	1	26.99	110.676	.000
Scrambling * Group		Linear		8.809	6	1.468	6.021	.000
Error(Scrambling)		Linear		21.948	90	.244		
Scope			Linear	157.926	1	157.926	201.868	.000
Scope * Group			Linear	12.779	6	2.130	2.722	.018
Error(Scope)			Linear	70.409	90	.782		
Subj QP * Scrambling	Linear	Linear		.113	1	.113	.626	.431
Subj QP * Scr. * Group	Linear	Linear		4.149	6	.691	3.845	.002
Error(Subj QP*Scrambling)	Linear	Linear		16.187	90	.180		
Subj QP * Scope	Linear		Linear	.725	1	.725	2.615	.109
Subj QP * Scope * Group	Linear		Linear	2.261	6	.377	1.360	.240
Error(Subj QP*Scope)	Linear		Linear	24.947	90	.277		
Scrambling * Scope		Linear	Linear	52.453	1	52.453	265.506	.000
Scrambling * Scope * Group		Linear	Linear	9.648	6	1.608	8.139	.000
Error(Scrambling*Scope)		Linear	Linear	17.780	90	.198		
Subj QP * Scr. * Scope	Linear	Linear	Linear	.425	1	.425	2.176	.144
Subj QP * Scr. * Scope * Group	Linear	Linear	Linear	1.811	6	.302	1.545	.173
Error(Subj QP*Scr.*Scope)	Linear	Linear	Linear	17.590	90	.195		

* 'Scr.' = Scrambling

Table 5F.4.iv:
Within-subjects contrasts (Types 1c and 1d compared with Type 1b)

	Type 1c compared with Type 1b					Type 1d compared with Type 1b				
	Type III Sum of Squares	df	Mean Square	F	Sig.	Type III Sum of Squares	df	Mean Square	F	Sig.
JJ	32.258	1	32.258	28.242	.000	47.895	1	47.895	56.148	.000
EJ int	19.220	1	19.220	33.188	.000	13.433	1	13.433	62.242	.000
EJ adv	15.471	1	15.471	6.842	.031	.640	1	.640	2.286	.169
CJ int	4.250	1	4.250	2.763	.157	4.250	1	4.250	2.763	.157
CJ adv	22.404	1	22.404	28.561	.001	11.111	1	11.111	20.161	.002
KJ int	23.835	1	23.835	23.072	.000	11.001	1	11.001	22.966	.000
KJ adv	59.203	1	59.203	92.120	.000	18.371	1	18.371	18.075	.001

Table 5F.4.v:
Within-subjects contrasts (Types 3c and 3d compared with Type 3b)

	Type 3c compared with Type 3b					Type 3d compared with Type 3b				
	Type III Sum of Squares	df	Mean Square	F	Sig.	Type III Sum of Squares	df	Mean Square	F	Sig.
JJ	74.885	1	74.885	141.707	.000	63.190	1	63.190	120.879	.000
EJ int	4.961	1	4.961	11.380	.004	4.961	1	4.961	11.380	.004
EJ adv	18.778	1	18.778	22.149	.002	3.738	1	3.738	9.894	.014
CJ int	11.207	1	11.207	6.286	.054	11.760	1	11.760	13.125	.015
CJ adv	31.360	1	31.360	92.235	.000	10.240	1	10.240	12.800	.007
KJ int	46.208	1	46.208	109.307	.000	27.028	1	27.028	68.983	.000
KJ adv	52.267	1	52.267	112.000	.000	28.843	1	28.843	37.260	.000

Appendix 6: *Wh*-QP test

About this appendix

This appendix relates to the second part of the main experimental study (Phase 3): investigation of questions with a QP-subject and *wh*-object. It should be used in conjunction with Chapter 5, Section 5.3. The appendix comprises six sections:

1. Appendix 6A: test items and pictures for the Japanese version of the test;
2. Appendix 6B: test battery in Chinese;
3. Appendix 6C: test battery in Korean;
4. Appendix 6D: L2ers' answer sheet in English;
5. Appendix 6E: L2ers' answer sheet in Japanese;
6. Appendix 6F: tables of results.

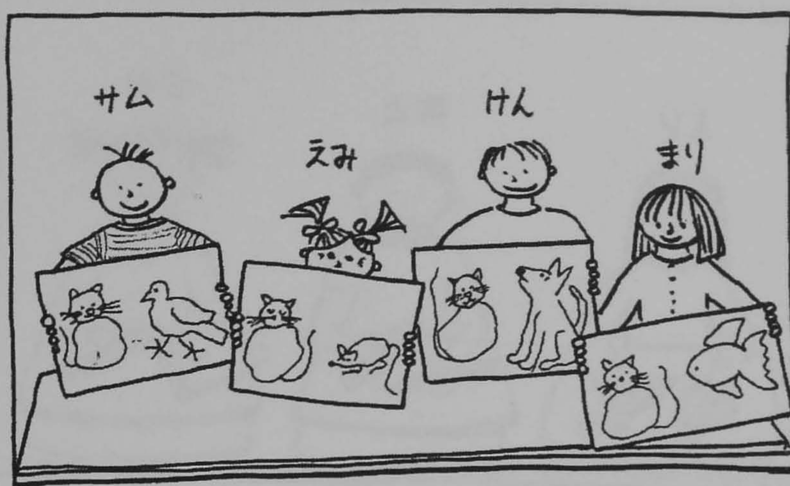
Note on Appendices 6A–6C:

The pictures used in the English, Chinese and Korean versions of the test were the same as those presented in Appendix 6A, except that names and text on the pictures appeared in the respective language of each test. The English test battery is not presented separately, because this version used the English translation sentences given in Appendix 6A (except that English names were used instead of Japanese names).

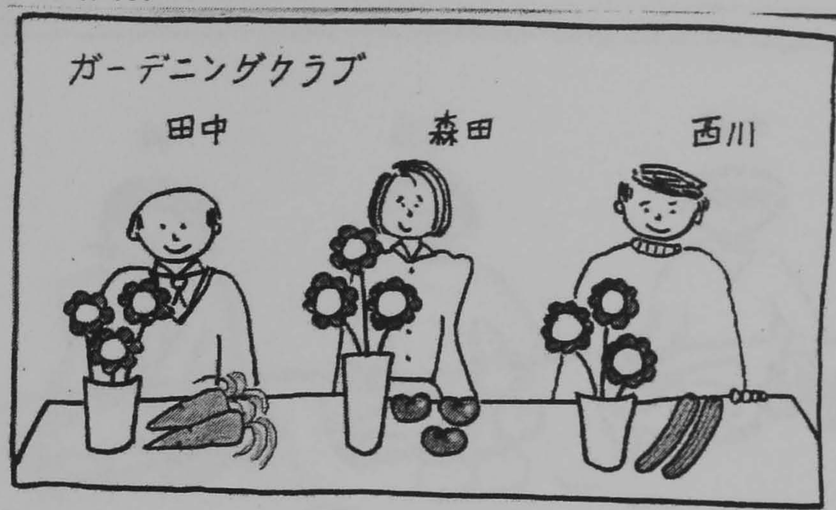
Appendix 6A: Test items and pictures in Japanese

Note: For convenience, English translations of the Japanese words appearing on the pictures are given below the pictures. These translations did not appear in the actual test.

index	**picture & Q/A**	**item no.**	
		order 1	**order 2**
	Picture for 1a & 1b: 		
1a	*Q:* 何をだれもが焼いたの？ Nani-o daremo-ga yaita no? what-ACC everyone-NOM bake.PAST Q 'What did everyone bake?' *A:* パンです。 Pan desu. bread COP 'Bread.'	1	20
1b	*Q:* (*As 1a*) *A:* Aさんはパンとケーキを、Bさんはパンとアップルパイを、Cさんはパンとクッキーを焼いた。 A-san-wa pan to keeki-o, B-san-wa pan to appuru-pai-o, C-san-wa A-*san*-TOP bread and cake-ACC, B-*san*-TOP bread and apple pie-ACC, C-*san*-TOP pan to kukkii-o yaita. bread and biscuit-ACC bake.PAST. 'A-*san* baked bread and a cake, B-*san* baked bread and a pie, and C-*san* baked bread and biscuits.'	13	8

index	picture & Q/A	item no. order 1	item no. order 2
	Picture for 2a & 2b:  **Translation (L-R):** Sam, Emi, Ken, Mari		
2a	*Q:* 何をだれもがかいたの？ Nani-o daremo-ga kaita no? what-ACC everyone-NOM draw.PAST Q 'What did everyone draw?' *A:* 猫です。 Neko desu. cat COP 'A cat.'	19	2
2b	*Q:* (*As 2a*) *A:* サム君は猫と鳥を、えみちゃんは猫とねずみを、けん君は猫と犬を、まりちゃんは猫と金魚をかいた。 Samu-kun-wa neko to tori-o, Emi-tyan-wa neko to nezumi-o, Ken-kun-wa Sam-*kun*-TOP cat and bird-ACC, Emi-*chan*-TOP cat and mouse-ACC, Ken-*kun*-TOP neko to inu-o, Mari-tyan-wa neko to kingyo-o kaita. and dog-ACC, Mari-*chan*-TOPcat and goldfish-ACC draw.PAST. 'Sam drew a cat and a bird, Emi drew a cat and a mouse, Ken drew a cat and a dog, and Mari drew a cat and a goldfish.'	6	15

index	picture & Q/A	item no. order 1	 order 2
	Picture for 3a & 3b:  **Translation (L-R):** Yuki, Taro, Rie		
3a	*Q:* 何をだれもが持ってきたの？ Nani-o daremo-ga motte-kita no? what-ACC everyone-NOM have-come.PAST Q 'What did everyone bring?' *A:* えんぴつです。 Enpitu desu. pencil COP 'A pencil.'	15	6
3b	*Q:* (*As 3a*) *A:* ゆきちゃんはえんぴつと本を、太郎君はえんぴつとペンを、 りえちゃんはえんぴつとはさみを持ってきた。 Yuki-tyan-wa enpitu to hon-o, Taroo-kun-wa enpitu to pen-o, Rie- Yuki-*chan*-TOP pencil and book-ACC, Taro-*kun*-TOP pencil and pen-ACC, Rie- tyan-wa enpitu to hasami-o motte-kita. *chan*-TOP pencil and scissors-ACC have-come.PAST. 'Yuki brought a pencil and a book, Taro brought a pencil and a pen, and Rie brought a pencil and a pair of scissors.'	4	17

index	picture & Q/A	item no.	
		order 1	order 2
	Picture for 4a & 4b:  **Translation (Line 1):** Gardening Club **(Line 2, L–R):** Tanaka, Morita, Nishikawa		
4a	*Q:* 何をだれもが作ったの？ Nani-o daremo-ga tukutta no? what-ACC everyone-NOM grow.PAST Q 'What did everyone grow?' *A:* 花です。 Hana desu. flower COP 'Flowers.'	11	10
4b	*Q:* (*As 4a*) *A:* 田中さんは花とにんじんを、森田さんは花とトマトを、西川さんは花ときゅうりを作った。 Tanaka-san-wa hana to ninzin-o, Morita-san-wa hana to tomato-o, Tanaka-*san*-TOP flower and carrot-ACC, Morita-*san*-TOP flower and tomato-ACC, Nisikawa-san-wa hana to kyuuri-o tukutta. Nishikawa-*san*-TOP flower and cucumber-ACC grow.PAST. 'Tanaka-*san* grew flowers and carrots, Morita-*san* grew flowers and tomatoes, and Nishikawa-*san* grew flowers and cucumbers.'	17	4

<table>
<tr><th rowspan="2">index</th><th rowspan="2">picture & Q/A</th><th colspan="2">item no.</th></tr>
<tr><th>order 1</th><th>order 2</th></tr>
<tr><td></td><td>Picture for 5a & 5b:

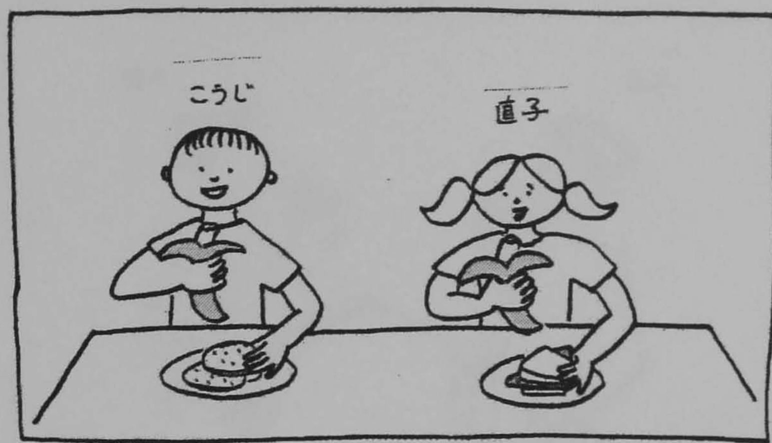
Translation (Top): At the supermarket
(Names, L–R): Hayashi, Yamada, Morimoto</td><td></td><td></td></tr>
<tr><td>5a</td><td>Q: 何をだれもが買ったの？
Nani-o daremo-ga katta no?
what-ACC everyone-NOM buy.PAST Q
‘What did everyone buy?’

A: 魚です。
Sakana desu.
fish COP
‘Fish.’</td><td>10</td><td>11</td></tr>
<tr><td>5b</td><td>Q: (As 5a)

A: 林さんは魚とバナナを、山田さんは魚とりんごを、森本さんは魚とたまねぎを買った。
Hayasi-san-wa sakana to banana-o, Yamada-san-wa sakana to ringo-o,
Hayashi-san-TOP fish and banana-ACC, Yamada-san-TOP fish and apple-ACC,

Morimoto-san-wa sakana to tamanegi-o katta.
Morimoto-san-TOP fish and onion-ACC buy.PAST.

‘Hayashi-san bought fish and bananas, Yamada-san bought fish and apples, and Morimoto-san bought fish and onions.’</td><td>2</td><td>19</td></tr>
</table>

index	picture & Q/A	item no. order 1	item no. order 2
	Picture for D1x & D1y:  **Translation (L-R):** Koji, Naoko		
D1x	*Q:* 何を女の子が食べたの？ Nani-o onnanoko-ga tabeta no? what-ACC girl-NOM eat.PAST Q 'What did the girl eat?' *A:* バナナとサンドイッチです。 Banana to sandoitti desu. banana and sandwich COP 'A banana and a sandwich.'	14	7
D1y	*Q:* 誰が何を食べたの？ Dare-ga nani-o tabeta no? who-NOM what-ACC eat.PAST Q 'Who ate what?' *A:* こうじ君はバナナとサンドイッチを、直子ちゃんはバナナとクッキーを食べた。 Koozi-kun-wa banana to sandoitti-o, Naoko-tyan-wa banana to Koji-*kun*-TOP banana and sandwich-ACC, Naoko-*chan*-TOP banana and kukkii-o tabeta. biscuit-ACC, eat.PAST. 'Koji ate a banana and a sandwich, and Naoko ate a banana and some biscuits.'	20	1

index	picture & Q/A	item no. order 1	item no. order 2
	Picture for D2x & D2y:  **Translation (L-R):** Tanaka, Suzuki		
D2x	*Q:* 誰が何を飲んだの？ Dare-ga nani-o nonda no? who-NOM what-ACC drink.PAST Q 'Who drank what?' *A:* 田中さんはオレンジジュースを、鈴木さんはビールを飲んだ。 Tanaka-san-wa orenzi-zyuusu-o, Suzuki-san-wa biiru-o nonda. Tanaka-*san*-TOP orange juice-ACC, Suzuki-*san*-TOP beer-ACC drink.PAST. 'Tanaka-*san* drank orange juice and Suzuki-*san* drank beer.'	12	9
D2y	*Q:* (*As D2x*) *A:* 田中さんはオレンジジュースを飲んだ。 Tanaka-san-wa orenzi-zyuusu-o nonda. Tanaka-*san*-TOP orange juice-ACC drink.PAST. 'Tanaka-*san* drank orange juice.'	18	3

<table>
<tr><th rowspan="2">index</th><th rowspan="2">picture & Q/A</th><th colspan="2">item no.</th></tr>
<tr><th>order 1</th><th>order 2</th></tr>
<tr><td></td><td>Picture for D3x & D3y:


Translation (L-R): Yumi, Yoshio, Shunji</td><td></td><td></td></tr>
<tr><td>D3x</td><td>Q: 誰が何をもらったの？
Dare-ga nani-o moratta no?
who-NOM what-ACC receive.PAST Q
‘Who got what?’

A: ゆみちゃんは手袋とマフラーを、よしお君は手袋と靴下を、しゅんじ君は手袋と帽子をもらった。
Yumi-tyan-wa tebukuro to mahuraa-o, Yosio-kun-wa tebukuro to
Yumi-chan-TOP gloves and scarf-ACC, Yoshio-kun-TOP gloves and

kutusita-o, Syunzi-kun-wa tebukuro to boosi-o moratta.
socks-ACC, Shunji-kun-TOP gloves and hat-ACC receive.PAST.

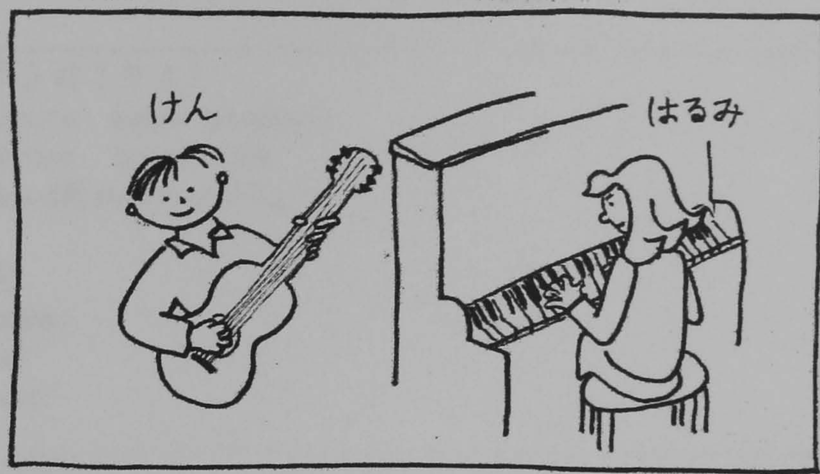
‘Yumi got gloves and a scarf, Yoshio got gloves and socks, and Shunji got gloves and a hat.’</td><td>8</td><td>13</td></tr>
<tr><td>D3y</td><td>Q: (As D3x)

A: 男の子は手袋をもらった。
Otokonoko-wa tebukuro-o moratta.
Boy-TOP gloves-ACC receive.PAST.
The boys got some gloves.</td><td>5</td><td>16</td></tr>
</table>

index	picture & Q/A	item no.	
		order 1	order 2
	Picture for D4x & D4y:  **Translation (L-R):** Mariko, Ayako, Satomi		
D4x	*Q:* 誰が何を買ったの？ Dare-ga nani-o katta no? who-NOM what-ACC buy.PAST Q 'Who bought what?' *A:* まり子さんは帽子を、あや子さんはかばんを、里美さんは靴を買った。 Mariko-san-wa boosi-o, Ayako-san-wa kaban-o, Satomi-san-wa Mariko-*san*-TOP hat-ACC, Ayako-*san*-TOP bag-ACC, Satomi-*san*-TOP kutu-o katta. shoe-ACC buy.PAST. 'Mariko bought a hat, Ayako bought a bag, and Satomi bought some shoes.'	7	14
D4y	*Q:* (*As D4x*) *A:* だれもが帽子を買った。 Daremo-ga boosi-o katta Everyone-NOM hat-ACC buy.PAST. 'Everyone bought a hat.'	16	5

<table>
<tr><th rowspan="2">index</th><th rowspan="2">picture & Q/A</th><th colspan="2">item no.</th></tr>
<tr><th>order 1</th><th>order 2</th></tr>
<tr><td></td><td>

Picture for D5x & D5y:



Translation (L-R): Ken, Harumi

</td><td></td><td></td></tr>
<tr><td>D5x</td><td>

Q: 何を男の子が弾いたの？
Nani-o otokonoko-ga hiita no?
what-ACC boy-NOM play.PAST Q
'Who played what?'

A: けんさんはギターを、はるみさんはピアノを弾いた。
Ken-san-wa gitaa-o, Harumi-san-wa piano-o hiita.
Ken-*san*-TOP guitar-ACC, Harumi-*san*-TOP piano-ACC play.PAST.
'Ken played the guitar and Harumi played the piano.'

</td><td>3</td><td>18</td></tr>
<tr><td>D5y</td><td>

Q: 誰が何を弾いたの？
Dare-ga nani-o hiita no?
who-NOM what-ACC play.PAST Q
'Who played what?'

A: はるみさんはピアノを弾いた。
Harumi-san-wa piano-o hiita.
Harumi-*san*-TOP piano-ACC play.PAST.
'Harumi played the piano.'

</td><td>9</td><td>12</td></tr>
</table>

Appendix 6B: *Wh*-QP test items in Chinese and English

Note: See Appendix 6A for (i) the pictures that accompanied each test item (the text on the pictures was in Chinese in the Chinese version); (ii) the item numbers.

index	Q/A
1a	*Q:* 每个人烤了什么？ Meigeren kaole shenme? everyone baked what 'What did everyone bake?' *A:* 面包。 Mianbao. bread 'Bread.'
1b	*Q:* (*As 1a*) *A:* A烤了面包和一块蛋糕，B烤了饼，C烤了面包面包和一个和饼干。 *A* kaole mianbao he yikuai dangao, *B* kaole mianbaohe yige bing, *C* *A* baked bread and a cake, *B* baked bread and a pie *C* kaole mianbao he binggan. baked bread and biscuits '*A* baked bread and a cake, *B* baked bread and a pie, and *C* baked bread and biscuits.'
2a	*Q:* 每个人画了什么？ Meigeren huale shenme? everyone draw what 'What did everyone draw?' *A:* 一只猫。 Yizhi mao. 'A cat.'
2b	*Q:* (*As 2a*) *A:* 萨姆画了一只猫和一只鸟，凯特画了一只猫和一只老鼠，戴维画了一只猫和一只狗，珍妮画了一只猫和一条金鱼。 Sam huale yizhi mao he yizhi niao, Kate huale yizhi mao he yizhi laoshu, Sam draw a cat and a bird Kate draw a cat and a mouse David huale yizhi mao he yizhi gou, Jane huale yizhi mao he yitiao jinyu. David draw a cat and a dog Jane draw a cat and a goldfish 'Sam drew a cat and a bird, Kate drew a cat and a mouse, David drew a cat and a dog, and Jane drew a cat and a goldfish.'
3a	*Q:* 每个人带了什么？ Meigeren daile shenme? everyone brought what 'What did everyone bring?' *A:* 一支铅笔。 Yizhi qianbi. 'A pencil.'

<table>
<tr><td>3b</td><td>Q: (As 3a)

A: 露茜带了一支铅笔和一本书，迪姆带了一支铅笔和一支钢笔，艾米带了一支铅笔和一把剪刀。
Lucy daile yizhi qianbi he yiben shu, Tim daile yizhi qianbi
Lucy brought a pencil and a book Tim brought a pencil

he yizhi gangbi, Amy daile yizhi qianbi he yiba jiandao.
and a pen, Amy brought a pencil and a scissors

'Lucy brought a pencil and a book, Tim brought a pencil and a pen, and Amy brought a pencil and a pair of scissors.'</td></tr>
<tr><td>4a</td><td>Q: 每个人种了什么？
Meigeren zhongle shenme?
everyone grew what
'What did everyone grow?'

A: 花。
Hua.
'Flowers.'</td></tr>
<tr><td>4b</td><td>Q: (As 4a)

A: 琼斯先生种了花和胡萝卜，威尔森夫人种了花和西红柿，布朗先生种了花和黄瓜。
Jones xiansheng zhongle hua he huluobo, Wilson furen zhongle hua
Jones Mr grew flower and carrot Wilson Mrs grew flower

he xihongshi Brown xiansheng zhongle hua he huanggua.
and tomato, Brown Mr grew flower and cucumber

'Mr Jones grew flowers and carrots, Mrs Wilson grew flowers and tomatoes, and Mr Brown grew flowers and cucumbers.'</td></tr>
<tr><td>5a</td><td>Q: 每个人买了什么？
Meigeren maile shenme?
everyone bought what
'What did everyone buy?'

A: 鱼。
Yu.
'Fish.'</td></tr>
<tr><td>5b</td><td>Q: (As 5a)

A: 史密斯夫人买了鱼和香蕉，沃克先生买了鱼和苹果，格林夫人买了鱼和洋葱。
Smith furen maile yu he xiangjiao, Walker xiansheng maile yu he
Smith Mrs bought fish and banana Walker Mr bought fish and

pinguo, Green furen maile yu he yangcong.
apple Green Mrs bought fish and onion

'Mrs Smith bought fish and bananas, Mr Walker bought fish and apples, and Mrs Green bought fish and onions.'</td></tr>
</table>

D1x	*Q:* 女孩吃了什么？ nuhai chile shenme? girl ate what 'What did the girl eat?' *A:* 一只香蕉和一个三文治 yizhi xiangjiao he yige sanwenzhi a banana and a sandwich 'A banana and a sandwich.'
D1y	*Q:* 谁吃了什么？ shui chile shenme? who ate what 'Who ate what?' *A:* 贝尔吃了一只香蕉和一个三文治，克莱尔吃了一只香蕉和一些饼干。 Bill chile yizhi xiangjiao he yige sanwenzhi, Clare chile yizhi xiangjiao Bill ate a banana and a sandwich Clare ate a banana he yixie binggan and some biscuits 'Bill ate a banana and a sandwich, and Clare ate a banana and some biscuits.'
D2x	*Q:* 谁喝了什么？ shui hele shenme? who drank what 'Who drank what?' *A:* 菲尔喝了桔子汁，福来德喝了啤酒。 Phil hele juzizhi, Fred hele pijiu Phil drank orange juice Fred drank beer 'Fred drank orange juice and Phil drank beer.'
D2y	*Q:* (*As D2x*) *A:* 菲尔喝了桔子汁。 Phil hele juzizhi Phil drank orange juice 'Phil drank orange juice.'
D3x	*Q:* 谁有什么？ shui you shenme? who have what 'Who got what?' *A:* 珍妮有手套 和围巾，詹姆斯有手套和袜子，罗博有手套和一顶帽子。 Jane you shoutao he weijin, James you shoutao he wazi, Rob you Jane have glove and scarf, James have glove and socks Rob have shoutao he yiding maozi glove and a hat 'Jane got gloves and a scarf, Rob got gloves and socks, and James got gloves and a hat.'

D3y	*Q:*	(*As D3x*)
	A:	男孩们有一些手套。 Nanhaimen you yixie shoutao boys have some glove. The boys got some gloves.
D4x[224]	*Q:*	谁带了什么？ shui daile shenme? who brought what 'Who brought what?'
	A:	玛丽亚带了一顶帽子，利兹带了一个书包，萨拉带了一些鞋子。 Marie daile yiding maozi, Liz daile yige shubao, Sarah daile yixie xiezi Marie brought a hat, Liz brought a bag Sarah brought some shoe 'Marie brought a hat, Liz brought a bag, and Sarah brought some shoes.'
D4y	*Q:*	(*As D4x*)
	A:	每个人带了一顶帽子。 meigeren daile yiding maozi everyone brought a hat 'Everyone brought a hat.'
D5x	*Q:*	男孩演奏了什么？ nanhai yanzoule shenme? boy played what 'What did the boy play?'
	A:	保罗演奏了吉它，罗拉演奏了钢琴 Paul yanzoule jita, Laura yanzoule gangqin Paul played guitar Laura played piano 'Paul played the guitar and Laura played the piano.'
D5y	*Q:*	谁演奏了什么？ shui yanzoule shenme? who played what 'Who played what?'
	A:	罗拉演奏了钢琴。 Laura yanzoule gangqin Laura played piano 'Laura played the piano.'

224 A mistranslation occurred on this test item and item D4y. The verb should have been 'bought' as in the Japanese, Korean and English versions.

Appendix 6C: *Wh*-QP test items in Korean

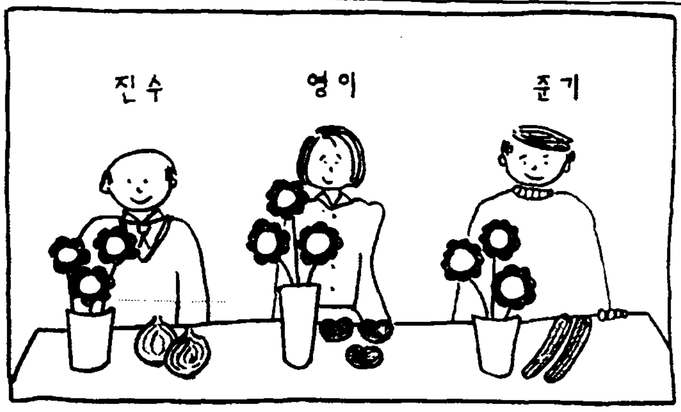
Notes:

i. See Appendix 6A for (i) the pictures that accompanied each test item (the text on the pictures was in Korean in the Korean version); (ii) the item numbers.

ii. Korean verbal suffixes (e.g., the past tense morpheme –*ss*, the complementiser -*ko*, and the declarative morpheme–*ta*) and the post-proper-name phonetic filler *i* are not glossed separately.

index	Q/A
1a	*Q:* 무엇을 누구나가 구웠니? Mwues-ul nwukwuna-ka kwuwess ni? what-ACC everyone-NOM bake.PAST Q 'What did everyone bake? *(scrambled)*' *A:* 빵을 구웠다. Ppang-ul kwuwessta. bread ACC baked '(Everyone) baked bread.'
1b	*Q:* (*As 1a*) *A:* 영이는 빵과 케이크을 구웠고, 수희는 빵과 파이를 구웠고, 진영이는 빵과 과자를 구웠다. Yengi-nun ppang kwa kheyikhu-ul kwuwessko, Swuhuy-nun ppang kwa phai-lul Yengi-TOP bread and cake-ACC baked, Swuhuy-TOP bread and pie-ACC kwuwessko, Cinyeng-i-nun ppang kwa kwaca-lul kwuwessta. baked, Cinyeng-TOP bread and biscuit-ACC baked 'Yengi baked bread and a cake, Swuhuy baked bread and a pie, and Cinyeng baked bread and biscuits.'
2a	*Q:* 무엇을 누구나가 그렸니? Mwues-ul nwukwuna-ka kulyess ni? what-ACC everyone-NOM draw PAST Q 'What did everyone draw? *(scrambled)*' *A:* 고양이를 그렸다. Koyangi-lul kulyessta. cat-ACC drew '(Everyone) drew a cat.'

2b	*Q:*	(*As 2a*)
	A:	경수는 고양이와 새를 그렸고, 수영이는 고양이와 생쥐를 그렸고, 희권이는 고양이와 개를 그렸고, 진경이는 고양이와 금붕어를 그렸다. Kyenswu-nun koyangi wa say-lul kulyessko, Swuyang-i-nun koyangi wa sayngcwi-lul Kyenswu-TOP cat and bird-ACC drew, Swuyang-TOP cat and mouse-ACC kulyessko, Huykwen-i-nun koyangi wa kay-lul kulyessko, Cinkyen-i-nun koyangi wa drew, Huykweni-TOP cat and dog-ACC drew, Cinkyen-TOP cat and kumpwunge-lul kulyessta. goldfish-ACC drew 'Kyenswu drew a cat and a bird, Swuyang drew a cat and a mouse, Huykwen drew a cat and a dog, and Cinkyen drew a cat and a goldfish.'
3a	*Q:*	무엇을 누구나가 가져왔니? Mwues-ul nwukwuna-ka kacyewass ni? what-ACC everyone-NOM bring PAST Q 'What did everyone bring? *(scrambled)*'
	A:	연필을 가져왔다. Yenphil-ul kacyewassta. pencil-ACC brought '(Everyone) brought a pencil.'
3b	*Q:*	(*As 3a*)
	A:	경희는 연필과 책을 가져왔고, 준영이는 연필과 펜을 가져왔고, 현수는 연필과 가위를 가져왔다. Kyenghuy-nun yenphil kwa chayk-ul kacyewassko, Cwunyeng-i-nun yenphil Kyenghuy-TOP pencil and book-ACC brought, Cwunyeng-TOP pencil kwa pheyn-ul kacyewassko, Hyenswu-nun yenphil kwa kawi-lul kacyewassta. and pen-ACC brought, Hyenswu-TOP pencil and scissors-ACC brought 'Kyenghuy brought a pencil and a book, Cwunyeng brought a pencil and a pen, and Hyenswu brought a pencil and a pair of scissors.'

4a & 4b[225]	
4a	*Q:* 무엇을 누구나가 키웠니? Mwues-ul nwukwuna-ka khiwess ni? what-ACC everyone-NOM grow.PAST Q 'What did everyone grow? *(scrambled)*' *A:* 꽃을 키웠다. Kkoch-ul khiwessta. flower-ACC grew '(Everyone) grew flowers.'
4b	*Q:* *(As 4a)* *A:* 진수는 꽃과 양파를 키우고, 영이는 꽃과 토마토를 키우고, 준기는 꽃과 오이를 키웠다. Cinswu-nun kkoch kwa yangpha-lul khiwuko, Yang-i-nun kkoch kwa thomatho-lul Cinswu-TOP flower and onion-ACC grow, Yang-TOP flower and tomato-ACC khiwuko, Cwunki-nun kkoch kwa oi-lul khiwessta. grow, Cwunki-TOP flower and cucumber-ACC grew 'Cinswu grew flowers and onions, Yang grew flowers and tomatoes, and Cwunki grew flowers and cucumbers.'
5a	*Q:* 무엇을 누구나가 샀니? Mwues-ul nwukwuna-ka sass ni? what-ACC everyone-NOM buy.PAST Q 'What did everyone buy? *(scrambled)*' *A:* 생선을 샀다. Sayngseng-ul sassta. fish-ACC bought '(Everyone) bought fish.'

[225] A separate picture was created for (4a) and (4b) in Korean, with the man on the left in the picture depicted with onions, instead of carrots as in the Japanese, Chinese and English versions. This is due to an error made while recording (4b) for the Korean audio tape.

	Q: (*As 5a*) *A:* 순이는 생선과 바나나를 샀고, 진수는 생선과 사과를 샀고, 영이는 생선과 양파를 샀다. Swun-i-nun sayngseng kwa panana-lul sassko, Cinswu-nun sayngseng kwa sakwa-lul Swun-TOP fish and banana-ACC bought, Cinswu-TOP fish and apple-ACC sassko, Yang-i-nun sayngseng kwa yangpha-lul sassta. bought, Yang-TOP fish and onion-ACC bought 'Swun bought fish and bananas, Cinswu bought fish and apples, and Yang bought fish and onions.'
D1x	*Q:* 무엇을 그 소녀가 먹었니? Mwues-ul ku sonye-ka mekess ni? what-ACC that girl-NOM eat.PAST Q 'What did the girl eat? *(scrambled)*' *A:* 바나나와 샌드위치를 먹었다. Panana wa sayntuwichi-lul mekessta. banana and sandwich-ACC ate '(She) ate a banana and a sandwich.'
D1y	*Q:* 누가 무엇을 먹었니? Nwuka mwues-ul mekess ni? who what-ACC eat.PAST Q 'Who ate what?' *A:* 철수는 바나나와 샌드위치를 먹었고, 영이는 바나나와 과자를 먹었다. Chelswu-nun panana wa sayngtuwichi-lul mekessko, Yang-i-nun panana wa Chelswu-TOP banana and sandwich-ACC ate, Yang-TOP banana and kwaca-lul mekessta. biscuit-ACC ate 'Chelswu ate a banana and a sandwich, and Yang ate a banana and some biscuits.'
D2x	*Q:* 누가 무엇을 마셨니? Nwuka mwues-ul masyess ni? who what-ACC drink.PAST Q 'Who drank what?' *A:* 준영이는 오렌지 주스를 마셨고 준기는 맥주를 마셨다. Cwunyang-nun oreynci cwusu-lul masyessko, Cwunki-nun meykcwu-lul masyessta. Cwunyang-TOP orange juice-ACC drank, Cwunki-TOP beer-ACC drank 'Cwunyang drank orange juice and Cwunki drank beer.'

D2y[226]	*Q:*	*(As D2x)*
	A:	철수가 오렌지 주스를 마셨다. Chelswu nun oreynci cwusu-lul masyessko Chelswu-TOP orange juice-ACC drank 'Chelswu drank orange juice.'
D3x	*Q:*	누가 무엇을 받았니? Nwuka mwues-ul patass ni? who what-ACC get.PAST Q 'Who got what?'
	A:	수영이는 장갑과 목도리를 받았고, 준수는 장갑과 양말을 받았고, 희권이는 장갑과 모자를 받았다. Swuyang-nun cangkap kwa moktori-lul patassko, Cwunswu-nun cangkap kwa Swuyang-TOP gloves and scarf-ACC got, Cwunswu-TOP gloves and yangmal-lul patassko, Huykwen-i-nun cangkap kwa moca-lul patassta. sock-ACC got, Huykwen-TOP gloves and hat-ACC got 'Swuyang got gloves and a scarf, Cwunswu got gloves and socks, and Huykwen got gloves and a hat.'
D3y	*Q:*	*(As D3x)*
	A:	소년들이 장갑을 받았다. Sonyentul-i cangkap-ul patassta. boys-NOM gloves-ACC got 'The boys got some gloves.'
D4x[227]	*Q:*	누가 무엇을 샀니? Nwuka mwues-ul sass ni? who what-ACC buy.PAST Q 'Who bought what?'
	A:	순이는 모자를 샀고, 영이는 가방을 샀고, 수희는 양말 몇 켤레를 샀다. Swun-i-nun moca-lul sassko, Yang-i-nun kapang-ul sass-ko, Swuhuy-nun Swun-TOP hat-ACC bought, Yang-TOP bag-ACC buy- bought, Swuhuy-TOP yangmal myech kkyelley-lul sass-ta sock some pair-ACC bought. 'Swun bought a hat, Yang bought a bag, and Swuhuy bought some socks.'

[226] Due to a translation error, the name in the answer for D2y was not one of the names on the picture. Therefore, the answer was incorrect for this reason in Korean. (In the other languages the answer was also expected to be judged incorrect, but on pragmatic grounds. See Chapter 5.)

[227] A mistranslation occurred on this test item. The third person in the picture (Swuhuy) was depicted with a pair of shoes but the Korean answer states she bought socks. Thus, the 'correct' answer on this item in Korean was '–1' or '–2' on the rating scale indicating that the answer is not possible. This was taken into account when examining the Korean distractor results.

D4y	*Q:*	(*As D4x*)
	A:	누구나 모자를 샀다. Nwukwuna moca-lul sassta. everyone hat-ACC bought 'Everyone bought a hat.'
D5x	*Q:*	무엇을 그 소년이 연주했니? Mwues-ul ku sonyen-i yencwu hayss ni? what-ACC that boy-NOM play.PAST Q 'What did the boy play?'
	A:	영수는 기타를 연주했고, 순이는 피아노를 연주했다. Yangswu-nun kitha-lul yencwu hayssko, Swun-i-nun phiano-lul yencwu hayssta. Yangswu-TOP guitar-ACC played, Swun-TOP piano-ACC played 'Yangswu played the guitar and Swun played the piano.'
D5y[228]	*Q:*	누가 무엇을 연주했니? Nwuka mwues-ul yencwu hayss ni? who what-ACC play.PAST Q 'Who played what?'
	A:	영이가 피아노를 연주했다. Yang-i-ka phiano-lul yencwu hayssta. Yang-NOM piano-ACC played 'Yang played the piano.'

[228] As for item D2y (see footnote 226), due to a translation error, the name in the answer for D5y was not one of the names on the picture. Therefore, the answer was incorrect for this reason in Korean. (In the other languages the answer was also expected to be judged incorrect, but on pragmatic grounds. See Chapter 5.)

Appendix 6D: L2ers' answer sheet in English

Instructions (Task 2)
For each test item you will see a picture. Underneath the picture, a question and answer will be shown. Please judge whether the answer is possible in the context of the picture. Indicate your answer by circling one of the options on the scale on your answersheet. The scale is as follows:

No, definitely not			Yes, perfectly		Can't decide
−2	−1	+1	+2		X

NOTE:
All the Japanese sentences are grammatically correct. This is not a 'find the mistakes' task.

Examples

(Is the answer possible?)

	No, definitely not			Yes, perfectly		Can't decide
Ex. 1	−2	−1	+1	+2		X
Ex. 2	−2	−1	+1	+2		X
Ex. 3	−2	−1	+1	+2		X
Ex. 4	−2	−1	+1	+2		X
Ex. 5	−2	−1	+1	+2		X
Ex. 6	−2	−1	+1	+2		X

****

Post-task comments (Please circle your answer, where appropriate)

1. Was the time for each test item: `too long` `just right` `too short`
2. Did you find any of the pictures difficult to understand?

   ....

   ....

3. Any further comments?

   ....

   ....

(Is the answer possible?)

	No, definitely not			Yes, perfectly		Can't decide
1	−2	−1	+1	+2		X
2	−2	−1	+1	+2		X
3	−2	−1	+1	+2		X
4	−2	−1	+1	+2		X
5	−2	−1	+1	+2		X
6	−2	−1	+1	+2		X
7	−2	−1	+1	+2		X
8	−2	−1	+1	+2		X
9	−2	−1	+1	+2		X
10	−2	−1	+1	+2		X
11	−2	−1	+1	+2		X
12	−2	−1	+1	+2		X
13	−2	−1	+1	+2		X
14	−2	−1	+1	+2		X
15	−2	−1	+1	+2		X
16	−2	−1	+1	+2		X
17	−2	−1	+1	+2		X
18	−2	−1	+1	+2		X
19	−2	−1	+1	+2		X
20	−2	−1	+1	+2		X

2

Appendix 6E: L2ers' answer sheet in Japanese

テストの方法（タスク２）
このテストはみなさんに絵と、その絵についての質問と答えを見ていただいて、その答えが、質問に、可能かどうかを判断していただくものです。判断は用紙のスケールを使い当てはまる数字を○で囲んで下さい。スケールは次の通りです：

（質問に、このように答えることが可能ですか）

いいえ、不可能です			はい、完全に可能です		分からない
−2	−1	+1	+2		X

注：テストの文はどれも文法的に正しいです。（つまり、これは「間違いを探す」ようなテストではありません。）

例：
（質問に、このように答えることが可能ですか）

	いいえ、不可能です			はい、完全に可能です		分からない
Ex. 1	−2	−1	+1	+2		X
Ex. 2	−2	−1	+1	+2		X
Ex. 3	−2	−1	+1	+2		X
Ex. 4	−2	−1	+1	+2		X
Ex. 5	−2	−1	+1	+2		X
Ex. 6	−2	−1	+1	+2		X

****

テスト後のコメント

1. **各件の時間は：　　長過ぎた　　ちょうどよかった　　短過ぎた**
2. **分かりにくい絵がありましたか。（どれか覚えていますか。）**

   ....

   ....
3. **その他のコメントはありますか。**

   ....

   ....

1

（質問に、このように答えることが可能ですか）

	いいえ、不可能です			はい、完全に可能です		分からない
1	–2	–1	+1	+2		X
2	–2	–1	+1	+2		X
3	–2	–1	+1	+2		X
4	–2	–1	+1	+2		X
5	–2	–1	+1	+2		X
6	–2	–1	+1	+2		X
7	–2	–1	+1	+2		X
8	–2	–1	+1	+2		X
9	–2	–1	+1	+2		X
10	–2	–1	+1	+2		X
11	–2	–1	+1	+2		X
12	–2	–1	+1	+2		X
13	–2	–1	+1	+2		X
14	–2	–1	+1	+2		X
15	–2	–1	+1	+2		X
16	–2	–1	+1	+2		X
17	–2	–1	+1	+2		X
18	–2	–1	+1	+2		X
19	–2	–1	+1	+2		X
20	–2	–1	+1	+2		X

2

Appendix 6F: Tables of results

6F.1. Raw data

Table 6F.1.i: Raw data, native control groups[229]

	id	order	mean: Wh-QPa	mean: Wh-QPb
Native Japanese	JJ01	2	2.8	1
	JJ02	2	3	0
	JJ03	2	3	0
	JJ04	2	2.6	0
	JJ05	2	3	0
	JJ06	2	2.2	3
	JJ07	2	0.6	0.6
	JJ08	2	2.6	1
	JJ09	2	3	1.2
	JJ10	2	2.4	2.8
	JJ11	2	2.2	2.6
	JJ12	2	3	1
	JJ13	2	3	1.4
	JJ14	2	3	3
	JJ15	2	3	3
	JJ16	2	2.8	0
	JJ17	2	3	0.8
	JJ18	2	2.6	1.8
Native English	EE01	1	2	3
	EE02	1	2.2	3
	EE03	1	2	3
	EE04	1	2	3
	EE05	1	1.6	3
	EE06	1	2	3
	EE07	1	2	3
	EE08	1	2	3
	EE09	1	2	3
	EE10	2	2	2.6
	EE11	2	2	3
	EE12	2	2	3
	EE13	2	2	3
	EE14	2	2	3
	EE15	2	2.6	3
	EE25	1	2.8	3
	EE36	2	1.4	3
	EE37	2	1.6	3
	EE38	2	2.4	3
	EE39	1	2	3
	EE40	1	1.6	3

[229] See Appendix 3 for age and gender information about participants.

	id	order	mean: Wh-QPa	mean: Wh-QPb
Native Chinese	CC01	1	1.6	3
	CC02	1	2.2	1.8
	CC03	1	1.2	2.6
	CC04	1	0.8	2.6
	CC05	1	1	3
	CC06	1	1.2	3
	CC07	1	1.2	2.2
	CC08	1	1.2	3
	CC09	1	1.2	2.4
	CC10	1	1	2.8
	CC11	1	2	2.4
	CC12	1	0.6	2.8
	CC13	1	1.4	2.4
	CC14	1	1.2	2
Native Korean	KK01	1	2.4	3
	KK02	1	2.25	2.6
	KK03	1	2.25	3
	KK04	1	3	2.4
	KK05	1	3	1.8
	KK07	1	3	3
	KK08	1	2.2	2.4
	KK09	1	2.2	2.4
	KK10	1	3	0
	KK11	1	2	1.8
	KK12	1	2.6	2
	KK13	1	3	1
	KK14	1	2	3
	KK15	1	1.8	3
	KK16	1	1	2.2
	KK17	1	2.25	2
	KK18	1	2.4	1.2
	KK19	1	2.4	0.6
	KK20	1	2.4	1.4
	KK21	1	1.8	2.4
	KK22	1	1.2	2.2
	KK23	1	1.2	2
	KK24	1	2.4	2
	KK25	1	2.2	2.8
	KK26	1	0.8	2.4

Table 6F.1.ii: Raw data, learners of Japanese[230]

	id	order	mean: Wh-QPa	mean: Wh-QPb
Intermediate EJ	EJ13	1	2.4	1.8
	EJ28	2	2.8	2.75
	EJ30	2	2.6	3
	EJ17	2	2	3
	EJ03	1	2.8	2.8
	EJ16	2	3	3
	EJ19	2	2.2	2.6
	EJ25	2	3	1
	EJ02	1	1	3
	EJ08	1	2.4	1.4
	EJ09	1	2.6	2.6
	EJ14	1	2.6	2.8
	EJ21	2	2.4	3
	EJ26	2	2.4	2.8
	EJ27	2	2.8	1
	EJ29	2	3	3
	EJ22	2	2.4	3
	EJ11	1	2.2	3
	EJ07	1	3	2.2
Advanced EJ	EJ05	1	2.8	3
	EJ23	2	2.2	3
	EJ24	2	3	1.4
	EJ06	1	3	2.4
	EJ12	1	2.8	1
	EJ04	1	2.4	0.6
	EJ20	2	3	3
	EJ15	2	1.8	3
	EJ10	1	3	0
Intermeidate CJ	CJ10	1	1.8	2.8
	CJ12	1	0.6	2.8
	CJ11	1	1.6	2.75
	CJ14	1	0.4	3
	CJ13	1	1.8	1.8
	CJ01	2	3	0
	CJ04	2	3	1.8
Avanced CJ	CJ03	2	3	2.6
	CJ15	2	2.2	3
	CJ09	2	2	2.8
	CJ18	2	3	0.6
	Cj08	2	3	0.6
	CJ16	2	1.8	2.6
	CJ07	2	3	3
	CJ02	2	3	0
	CJ06	2	3	2.2
	CJ05	2	3	0.2

[230] See Appendix 3 for proficiency task scores and age and gender data.

	id	order	mean: Wh-QPa	mean: Wh-QPb
Intermediate KJ	KJ19	1	1.6	3
	KJ14	1	3	2.8
	KJ29	1	2.4	2.2
	KJ39	1	1	2.4
	KJ25	1	2	3
	KJ35	1	2.8	0.8
	KJ04	1	2.4	2.8
	KJ18	1	0.2	2.4
	KJ20	1	2.4	0.8
	KJ21	1	2.5	3
	KJ26	1	3	3
	KJ42	1	2.6	2
	KJ05	1	2.8	1.6
	KJ32	1	1.8	1.8
	KJ33	1	3	1.8
	KJ15	1	2.8	0
	KJ27	1	2.8	2.8
	KJ30	1	1.6	3
	KJ31	1	1	3
	KJ41	1	3	3
	KJ11	2	1.4	2.2
	KJ17	1	0.8	2.4
Advanced KJ	KJ23	1	1.2	2.8
	KJ36	1	3	2.2
	KJ06	1	2.8	2.6
	KJ13	2	0.8	3
	KJ02	2	3	0.8
	KJ12	2	1.8	3
	KJ08	2	2.4	1
	KJ09	2	2.8	1.6
	KJ22	1	2	2.8
	KJ38	1	1.2	1.8
	KJ01	2	3	1
	KJ10	2	3	0
	KJ34	1	2.8	0.4
	KJ03	2	3	0
	KJ07	1	3	3

Table 6F.1.iii: Mean ratings on *Wh*-QP test tokens by native Japanese, English, Chinese and Korean groups

index	JJ		EE		CC		KK	
	mean	SD	mean	SD	mean	SD	mean	SD
1a	2.72	0.57	1.81	0.6	1.29	0.91	1.5	1.19
2a	2.5	0.79	2	0.45	1.14	0.66	2.6	0.65
3a	2.5	0.79	2.05	0.59	1.07	0.62	2.48	0.82
4a	2.78	0.55	2.1	0.44	1.5	0.65	2.24	0.93
5a	2.78	0.55	2.1	0.44	1.36	0.84	1.92	1.15
1b	1.17	1.1	3	0	2.36	0.74	2.04	1.21
2b	1.39	1.2	2.9	0.44	2.93	0.27	2.33	0.92
3b	1.33	1.24	3	0	2.36	0.84	2.3	0.88
4b	1.22	1.26	3	0	2.43	1.02	1.42	1.14
5b	1.33	1.24	3	0	2.79	0.58	2.48	0.9

Table 6F.1.iv: Mean ratings on *Wh*-QP test tokens by L2 groups

index	EJ int		EJ adv		CJ int		CJ adv		KJ int		KJ adv	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
1a	2.16	1.01	2.22	1.09	2.29	0.95	2.9	0.32	1.77	1.11	2.07	1.22
2a	2.42	0.9	2.67	0.71	2	1.41	2.3	1.16	2.41	0.8	2.4	0.99
3a	2.63	0.6	2.78	0.67	1.57	1.4	2.6	0.7	2.19	1.12	2.73	0.8
4a	2.68	0.58	2.78	0.44	1.43	1.4	2.8	0.42	2.27	0.94	2.33	1.11
5a	2.63	0.6	2.89	0.33	1.43	1.4	2.9	0.32	2	1.23	2.4	1.12
1b	2.39	0.98	2.11	1.17	1.86	1.35	1.8	1.4	1.91	1.34	1.47	1.3
2b	2.53	0.7	1.78	1.3	2.29	1.25	1.7	1.49	2.32	1.13	1.87	1.3
3b	2.53	0.7	1.89	1.36	2.17	1.33	1.9	1.37	2.64	0.85	1.87	1.3
4b	2.42	0.96	2	1.32	2	1.41	1.7	1.34	1.82	1.33	1.67	1.23
5b	2.68	0.67	1.89	1.36	2.29	1.25	1.7	1.49	2.64	0.9	1.8	1.37

6F.2: Repeated measures ANOVA[231]

Table 6F.2.i: Descriptive statistics

	GROUP	Mean	Std. Deviation	N
WHQPA	JJ	2.6556	.58533	18
	EE	2.0095	.31923	21
	CC	1.2714	.42685	14
	KK	2.1900	.62467	25
	EJ int	2.5053	.47314	19
	EJ adv	2.6667	.43589	9
	CJ int	1.7429	1.02446	7
	CJ adv	2.7000	.49216	10
	KJ int	2.1318	.83574	22
	KJ adv	2.3867	.78364	15
WHQPB	JJ	1.2889	1.14629	18
	EE	2.9810	.08729	21
	CC	2.5714	.39111	14
	KK	2.1040	.78128	25
	EJ int	2.5132	.69279	19
	EJ adv	1.9333	1.19583	9
	CJ int	2.1357	1.06408	7
	CJ adv	1.7600	1.24651	10
	KJ int	2.2636	.84715	22
	KJ adv	1.7333	1.12800	15

Table 6F.2.ii: Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	ANSWER	Type III Sum of Squares	df	Mean Square	F	Sig.
ANSWER	Linear	.649	1	.649	.870	.352
ANSWER * GROUP	Linear	49.205	9	5.467	7.335	.000
Error(ANSWER)	Linear	111.809	150	.745		

[231] Sphericity is assumed, as Mauchly's test of sphericity was non-signficant. Equality of variance is not assumed, as Levene's test was significant:

Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
WHQPA	3.998	9	150	.000
WHQPB	9.416	9	150	.000

Table 6F.2.iii: Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1292.811	1	1292.811	3346.469	.000
GROUP	13.213	9	1.468	3.800	.000
Error	57.948	150	.386		

6F.3: Within-groups comparison (t-tests)[232]

		Paired comparisons							
					95% confidence interval of the difference				
		Mean	SD	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
JJ	WHQPA - WHQPB	1.3667	1.32887	.31322	.7058	2.0275	4.363	17	.000
EJ int	WHQPA - WHQPB	-.0079	.94638	.21711	-.4640	.4482	-.036	18	.971
EJ adv	WHQPA - WHQPB	.7333	1.40000	.46667	-.3428	1.8095	1.571	8	.155
KJ int	WHQPA - WHQPB	-.1318	1.32139	.28172	-.7177	.4541	-.468	21	.645
KJ adv	WHQPA - WHQPB	.6533	1.69279	.43708	-.2841	1.5908	1.495	14	.157
CJ int	WHQPA - WHQPB	-.3929	1.96986	.74454	-2.2147	1.4290	-.528	6	.617
CJ adv	WHQPA - WHQPB	.9400	1.57212	.49715	-.1846	2.0646	1.891	9	.091
EE	WHQPA - WHQPB	-.9714	.33037	.07209	-1.1218	-.8210	-13.475	20	.000
CC	WHQPA - WHQPB	-1.3000	.70493	.18840	-1.7070	-.8930	-6.900	13	.000
KK	WHQPA - WHQPB	.0860	1.13363	.22673	-.3819	.5539	.379	24	.708

[232] The table gives two-tailed significance values, since this is the only option SPSS provides. The one-tailed significance values reported in Chapter 5 were obtained simply by dividing the two-tailed value by two. (One-tailed values are appropriate because the hypotheses tested are directional.)

Table 6F.4.i:
Between-groups post hoc multiple comparisons (Games-Howell): *Wh*-QPa

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
JJ	EJ int	.1503	.17555	.997	-.4448	.7454
	EJ adv	-.0111	.20036	1.000	-.7171	.6949
	KJ int	.5237	.22535	.398	-.2339	1.2814
	KJ adv	.2689	.24490	.980	-.5776	1.1154
	CJ int	.9127	.41105	.511	-.8344	2.6598
	CJ adv	-.0444	.20798	1.000	-.7747	.6859
	EE	.6460	.15455	.009	.1116	1.1805
	CC	1.3841	.17902	.000	.7733	1.9949
	KK	.4656	.18612	.300	-.1593	1.0904
EJ int	EJ adv	-.1614	.18136	.995	-.8161	.4933
	KJ int	.3734	.20864	.737	-.3320	1.0789
	KJ adv	.1186	.22961	1.000	-.6870	.9241
	CJ int	.7624	.40214	.675	-.9914	2.5162
	CJ adv	-.1947	.18975	.986	-.8761	.4866
	EE	.4957	.12898	.017	.0570	.9345
	CC	1.2338	.15747	.000	.6962	1.7715
	KK	.3153	.16550	.666	-.2373	.8678
EJ adv	KJ int	.5348	.22991	.405	-.2561	1.3258
	KJ adv	.2800	.24910	.976	-.5934	1.1534
	CJ int	.9238	.41357	.504	-.8262	2.6738
	CJ adv	-.0333	.21292	1.000	-.8023	.7357
	EE	.6571	.16113	.033	.0410	1.2732
	CC	1.3952	.18473	.000	.7278	2.0627
	KK	.4767	.19162	.330	-.2001	1.1535
KJ int	KJ adv	-.2548	.26961	.993	-1.1712	.6615
	CJ int	.3890	.42624	.991	-1.3554	2.1334
	CJ adv	-.5682	.23658	.362	-1.3803	.2439
	EE	.1223	.19131	1.000	-.5353	.7799
	CC	.8604	.21157	.009	.1433	1.5774
	KK	-.0582	.21762	1.000	-.7882	.6719
KJ adv	CJ int	.6438	.43689	.875	-1.1094	2.3970
	CJ adv	-.3133	.25527	.960	-1.2043	.5776
	EE	.3771	.21399	.748	-.3937	1.1480
	CC	1.1152	.23228	.003	.3008	1.9297
	KK	.1967	.23780	.997	-.6279	1.0212
CJ int	CJ adv	-.9571	.41732	.474	-2.7062	.7919
	EE	-.2667	.39343	.999	-2.0323	1.4990
	CC	.4714	.40367	.957	-1.2813	2.2242
	KK	-.4471	.40687	.970	-2.1962	1.3019
CJ adv	EE	.6905	.17051	.032	.0470	1.3340
	CC	1.4286	.19297	.000	.7354	2.1218
	KK	.5100	.19958	.296	-.1928	1.2128
EE	CC	-.7381	.13367	.000	-1.2056	-.2706
	KK	-.9186	.16918	.000	-1.4892	-.3480
CC	KK	.9186	.16918	.000	.3480	1.4892

Table 6F.4.ii:
Between-groups post hoc multiple comparisons (Games-Howell): *Wh*-QPb

GROUP	GROUP	Mean Difference	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
JJ	EJ int	-1.2243	.31346	.016	-2.3003	-.1482
	EJ adv	-.6444	.48155	.929	-2.4059	1.1170
	KJ int	-.9747	.32499	.122	-2.0816	.1321
	KJ adv	-.4444	.39727	.979	-1.7993	.9104
	CJ int	-.8468	.48451	.753	-2.7011	1.0075
	CJ adv	-.4711	.47789	.989	-2.1918	1.2495
	EE	-1.6921	.27085	.000	-2.6690	-.7151
	CC	-1.2825	.28970	.007	-2.2988	-.2663
	KK	-.8151	.31211	.258	-1.8853	.2551
EJ int	EJ adv	.5798	.42913	.919	-1.0968	2.2565
	KJ int	.2495	.24059	.988	-.5572	1.0562
	KJ adv	.7798	.33179	.398	-.3829	1.9425
	CJ int	.3774	.43245	.993	-1.4351	2.1900
	CJ adv	.7532	.42502	.740	-.8680	2.3743
	EE	-.4678	.16007	.168	-1.0397	.1041
	CC	-.0583	.19023	1.000	-.7083	.5917
	KK	.4092	.22288	.709	-.3361	1.1544
EJ adv	KJ int	-.3303	.43762	.998	-2.0149	1.3543
	KJ adv	.2000	.49367	1.000	-1.5948	1.9948
	CJ int	-.2024	.56625	1.000	-2.3137	1.9090
	CJ adv	.1733	.56059	1.000	-1.8527	2.1993
	EE	-1.0476	.39906	.333	-2.7155	.6203
	CC	-.6381	.41209	.844	-2.3051	1.0290
	KK	-.1707	.42814	1.000	-1.8455	1.5042
KJ int	KJ adv	.5303	.34270	.860	-.6588	1.7195
	CJ int	.1279	.44088	1.000	-1.6836	1.9394
	CJ adv	.5036	.43359	.967	-1.1284	2.1356
	EE	-.7173	.18161	.020	-1.3555	-.0791
	CC	-.3078	.20868	.892	-1.0168	.4012
	KK	.1596	.23882	1.000	-.6367	.9560
KJ adv	CJ int	-.4024	.49657	.997	-2.2833	1.4785
	CJ adv	-.0267	.49011	1.000	-1.7833	1.7299
	EE	-1.2476	.29187	.018	-2.3304	-.1649
	CC	-.8381	.30944	.242	-1.9513	.2751
	KK	-.3707	.33052	.977	-1.5284	.7871
CJ int	CJ adv	.3757	.56314	.999	-1.7097	2.4611
	EE	-.8452	.40263	.579	-2.6908	1.0003
	CC	-.4357	.41555	.976	-2.2601	1.3886
	KK	.0317	.43147	1.000	-1.7801	1.8435
CJ adv	EE	-1.2210	.39464	.180	-2.8204	.3785
	CC	-.8114	.40780	.622	-2.4162	.7933
	KK	-.3440	.42402	.997	-1.9630	1.2750
EE	CC	.4095	.10625	.040	.0142	.8048
	KK	.8770	.15741	.000	.3313	1.4226
CC	KK	.4674	.18800	.309	-.1651	1.1000

References

AJLT (Association for Japanese Language Teaching). 1996/1997. *Japanese For Busy People, Kana Version, Volumes 1–3*. Tokyo: Kodansha.

Aoun, Joseph & Norbert Hornstein. 1985. Quantifier types. *Linguistic Inquiry* 16: 623–36.

Aoun, Joseph & Yen-hui Audrey Li. 1989. Constituency and scope. *Linguistic Inquiry* 20: 141–72.

Aoun, Joseph & Yen-hui Audrey Li. 1993. *Syntax of scope*. Cambridge, MA: MIT Press.

Aoun, Joseph & Dominique Sportiche. 1983. On the formal theory of government. *Linguistic Review* 2: 211–236.

Beck, Sigrid & Shu-Sook Kim. 1997. On wh- and operator scope in Korean. *Journal of East Asian Linguistics* 6: 339–384.

Beghelli, Filippo. 1995. *The phrase structure of quantifier scope*. PhD dissertation, UCLA.

Beghelli, Filippo. 1997. The syntax of distributivity and pair-list readings. In Anna Szabolcsi (ed.) *Ways of scope taking*. Dordrecht: Kluwer. 349–408.

Beghelli, Filippo & Tim Stowell. 1997. Distributivity and negation: the syntax of each and every. In Anna Szabolcsi (ed.) *Ways of scope taking*. Dordrecht: Kluwer. 71–107.

Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory* 6: 291–352.

Bowring, Richard J. & Haruko U. Laurie. 1992. *An introduction to modern Japanese, Volumes 1–2*. Cambridge: Cambridge University Press.

Burzio, Luigi. 1986. *Italian syntax: A Government and Binding approach.* Dordrecht: Reidel.

Chen, Dongdong. 1996. *L2 acquisition of English psych predicates by native speakers of Chinese and French.* PhD dissertation: McGill University.

Chierchia, Gennaro. 1991. Functional WH and weak crossover. In Dawn Bates (ed.) *Proceedings of theTenth West Coast Conference on Formal Linguistics.* Stanford, CA: Center for the Study of Language and Information. 75–90.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1: 181–234.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris: Dordrecht.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.) *Principles and parameters in comparative grammar*. Cambridge, MA: MIT Press. 417–454.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (eds.) *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. 1–52.

Clark-Peasey, Sarah. 1999. *Semantic intuitions of Chinese learners of English, regarding entailments involving quantifiers*. MA dissertation, University of York.

Collick, Martin, David P. Dutcher, Munekazu Tanabe & Minoru Kaneko. 2002. *Kenkyusha's New College Japanese-English Dictionary, 5th Edition*. Tokyo: Kenkyusha.

Cowan, Ron. & Yukiko Abe Hatasa. 1994. Investigating the validity and reliability of native speaker and second-language learners judgments about sentences. In Tarone, Elaine E., Gass, Susan M., & Cohen, Andrew D. (eds.) *Research methodology in second-language acquisition*. Hillsdale, NJ: Lawrence Erlbaum. 287–302.

Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Dekydtspotter, Laurent, Rex A. Sprouse & Kimberley A. Swanson. 2001. Reflexes of mental architecture in second language acquisition: The interpretation of *combien* extractions in English-French interlanguage. *Language Acquisition* 9: 175–227.

Dekydtspotter, Laurent, Rex A. Sprouse, Kimberley A. Swanson & Rachel Thyre. 1999. Semantics, pragmatics and second language acquisition: A case of combien extractions. In Annabel Greenhill, Heather Littlefield & Cheryl Tano, (eds.) *Proceedings of the 23rd annual Boston University conference on language development, Volume 1*. Somerville, MA: Cascadilla Press. 162–71.

Dekydtspotter, Laurent, Rex A. Sprouse & Rachel Thyre. 1998. Evidence for Full Access in L2 acquisition from the interpretive interface: quantification at a distance in English-French interlanguage. In Annabel Greenhill, Mary Hughes, Heather Littlefield & Hugh Walsh (eds.) *Proceedings of the 22nd annual Boston University conference in language development, Volume 1.* Somerville MA: Cascadilla Press. 141–152.

Diesing, Molly. 1992. *Indefinites.* Cambridge, MA: MIT Press.

Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure.* Cambridge: Cambridge University Press.

Erteschik-Shir, Nomi. 1999. Focus structure and scope. In Georges Rebuschi & Laurence Tuller (eds.) *The grammar of focus.* Amsterdam: John Benjamins. 119–150.

Field, Andy. 2000. *Discovering statistics using SPSS for Windows.* London: Sage Publications.

Fodor, Janet D. 1982. The mental representation of quantifiers. In Stanley Peters & Esa Saarinen (eds.) *Processes, beliefs & and questions.* Dordrecht: Reidel. 129–164.

Gill, Kook-Hee. 2002. *Wh-quantification in Korean, Japanese and Malayalam: A comparative view.* Paper presented at workshop on disjunction, polarity, quantifiers and WH, 22–23 November, University of York.

Gill, Kook-Hee, Steve Harlow & George Tsoulas. 2002. *Disjunction, quantification and WH.* Manuscript, University of York.

Goodman, Nelson. 1951. *The structure of appearance.* Cambridge, MA: Harvard University Press.

Grimshaw, Jane. 1990. *Argument structure.* Cambridge, MA: MIT Press.

Guasti, Maria Teresa. 2002. *Language acquisition: the growth of grammar.* Cambridge, MA; MIT Press.

Haegeman, Liliane. 1985. Scope phenomena in English and Dutch and L2 acquisition: a case study. *Second Language Research* 1: 118–150.

Hagstrom, Paul. 1998. *Decomposing questions.* PhD dissertation, MIT.

Hasegawa, Nobuko. 1991. Affirmative polarity items and negation in Japanese. In Carol Georgopoulos & Roberta Ishihara (eds.) *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda.* Dordrecht: Kluwer. 271–285.

Hasegawa, Nobuko. 1993. Floating quantifiers and bare NP constructions. In Nobuko Hasegawa (ed.) *Japanese syntax in comparative grammar*. Tokyo: Kurosio. 115–145.

Hayashishita, J.-R. 1999. Two ways of deriving distributive readings. *Proceedings of the 23rd annual Penn Linguistics Colloquium: University of Pennsylvania Working Papers in Linguistics, Volume 6*. 201–216.

Hayashishita, J.-R. 2000a. More on two ways of deriving distributive readings. In Mineharu Nakayama & Charles J. Quinn, Jr. (eds.) *Japanese/Korean Linguistics, Volume 9*. Stanford, CA: CSLI Publications. 283–296.

Hayashishita, J.-R. 2000b. Scope ambiguity and 'scrambling'. In Roger Billerey & Brook Danielle Lillehaugen (eds.) *Proceedings of the Nineteenth West Coast Conference on Formal Linguistics*. 204–217.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA & Oxford: Blackwell.

Hirakawa, Makiko. 1999. L2 acquisition of Japanese unaccusative verbs by speakers of Englisih and Chinese. In Kazue Kanno (ed.) The *acquisition of Japanese as a second language*. Amsterdam/Philadelphia: John Benjamins. 89–113.

Hoji, Hajime. 1985. *Logical form constraints and configurational structures in Japanese*. PhD dissertation, University of Washington.

Hoji, Hajime. 1986. Scope interpretation in Japanese and its theoretical implications. *Proceedings of the West Coast Conference on Formal Linguistics, Volume 5*. Stanford: Stanford Linguistics Association, Stanford University. 87–101.

Hornstein, Norbert. 1995. *Logical Form: From GB to Minimalism*. Oxford: Blackwell.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. PhD dissertation, MIT.

Huang, C.-T. James. 1993. Reconstruction and the structure of VP: Some theoretical consequences. *Linguistic Inquiry* 24: 103–138.

Ioup, Georgette. 1975. Some universals for quantifier scope. In John Kimball (ed.) *Syntax and Semantics, Volume 4*. New York: Academic Press. 37–58.

Ishida, Toshiko. 1992. *Nyuumon: Nihongo tesuto hou* (An introduction to Japanese language testing). Tokyo: Taishukan.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, MA; MIT Press.

Jang, Youngjun. 1999. Two types of question and existential quantification. *Linguistics* 37: 847–869.

Jonz, Jon. 1990. Another turn in the conversation: What does cloze measure? *TESOL Quarterly* 24: 61–83.

Kempson, Ruth M. & Annabel Cormack. 1981. Ambiguity and quantification. *Linguistics and Philosophy* 4: 259–309.

Kawashima, Ruriko. 1994a. Deriving paradoxical flip-flop change in quantificational force in Japanese. In Masatoshi Koizumi & Hiroyuki Ura (eds.) *Formal approaches to Japanese linguistics 1: MIT Working Papers in Linguistics, Volume 24*. Cambridge, MA: Department of Linguistics, MIT. 121–140.

Kawashima, Ruriko. 1994b. *The structure of noun phrases and the interpretation of quantificational NPs in Japanese*. PhD dissertation, Cornell University.

Kawashima, Ruriko. 1998. The structure of extended nominal phrases: The scrambling of numerals, approximate numerals, and quantifiers in Japanese. *Journal of East Asian Studies* 7: 1–26.

Kim, Soo Won. 1989. The QP status of wh-phrases in Korean and Japanese. In E. Jane Fee (ed.) *Proceedings of the Eighth West Coast Conference on Formal Linguistics*. Stanford: CSLI Publications. 358–372.

Kim, Soo Won. 1991. *Chain scope and quantification structure*. PhD dissertation, Brandeis University, MA.

Kim, Young-Joo & Richard Larson. 1989. Scope interpretation and the syntax of psych-verbs. *Linguistic Inquiry* 20: 681–688.

Kitagawa, Yoshihisa. 1986. *Subjects in Japanese and English*. PhD dissertation, University of Massachusetts.

Kitahara, Hisatsugu. 1996. Raising quantifiers without quantifier raising. In Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson & C. Jan-Wouter Zwart (eds.) *Minimal ideas: Syntactic studies in the minimalist framework*. Amsterdam: John Benjamins. 189–198.

Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85: 211–58.

Kroch, Anthony S. 1975. *The semantics of scope in English*. PhD dissertation, MIT.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuno, Susumu. 1991. Remarks on quantifier scope. In Heizo Nakajima (ed.) *Current English linguistics in Japan*. Berlin: Mouton de Gruyter. 261–287.

Kuno, Susumu & Ken-ichi Takami. 2002. *Quantifier scope*. Tokyo: Kurosio.

Kuno, Susumu, Ken-ichi Takami & Yuru Wu. 1999. Quantifier scope in English, Chinese, and Japanese. *Language* 75: 63–111.

Kuno, Susumu, Ken-ichi Takami & Yuru Wu. 2000. Response to Aoun and Li. *Language* 76: 134–143.

Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*. PhD dissertation, MIT. Reprinted, New York: Garland Press. 1979.

Kuroda, S.-Y. 1970. Remarks on the notion of subject with reference to words like also, even and only, Part II. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo* 4: 127–152.

Kuroda, S.-Y. 1988. Whether we agree or not: A comparative syntax of English and Japanese. *Linguisticae Investigationes* 21: 1–46.

Kuroda, S.-Y. 1992. *Japanese syntax and semantics: collected papers*. Dordrecht: Kluwer.

Kurtzman, Howard S. & Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48: 243–279.

Lakoff, George. 1972. Linguistics and natural logic. In Donald Davidson & Gilbert Harman (eds.) *Semantics of natural language*. Dordrecht: Reidel. 545–665.

Larsen-Freeman, Diane & Michael H. Long. 1991. *An introduction to second language acquisition research*. London & New York: Longman.

Lasnik, Howard & Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235–289.

Lee, Chungmin. 1996. Negative polarity items in English and Korean. *Language Sciences* 18: 505–523.

Lee, Thomas Hun-tak, Virginia Yip & Wang Chuming. 1999a. Inverse scope in Chinese-English interlanguage. *Lingua Posnaniensis* 41: 39–56.

Lee, Thomas Hun-tak, Virginia Yip & Wang Chuming. 1999b. Rethinking isomorphism as a scope principle for Chinese and English. In Sun Chaofen (ed.) *Proceedings of the Tenth North American Chinese Linguistics Conference*. Graduate students in linguistics: Department of Linguistics, University of Southern California. 169–186.

Liu, Feng-Hsi. 1990. *Scope dependency in English and Chinese*. PhD dissertation: UCLA.

Liu, Feng-Hsi. 1997. *Scope and specificity*. Amsterdam: John Benjamins.

MacWhinney, Brian. 1987. The competition model. In Brian. MacWhinney (ed.) *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum. 249–308.

MacWhinney, Brian. & Elizabeth Bates. 1989. *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.

Marsden, Heather. 2003. Inverse scope in L2 Japanese. In Barbara Beachley, Amanda Brown & Frances Conlin (eds.) *Proceedings of the 27th annual Boston University Conference on Language Development, Volume 2*. Somerset, MA: Cascadilla Press. 496–507.

Marsden, Heather. 2004. L2 knowledge of quantifier scope in Korean and English learners of Japanese. In Marcela Cazzoli-Goeta, Makiko Mukai & Lieve Van Espen (eds.) *Durham working papers in linguistics, Volume 10*. Durham, UK: School of Linguistics and Language, University of Durham.

Martin, Samuel, E. 1992. *A reference grammar of Korean: A complete guide to the grammar and history of the Korean language*. Tokyo: Charles E. Tuttle.

Martohardjono, Gita & James W. Gair. 1993. Apparent UG accessibility in second language acquisition: Misapplied principles or principled misapplications? In Fred R. Eckman (ed.) *Confluence: Linguistics, L2 acquisition, speech pathology*. Amsterdam: John Benjamins. 79–103.

May, Robert. 1977. *The grammar of quantification*. PhD dissertation, MIT.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, MA: MIT Press.

McCawley, James D. 1977. Lexicographic notes on English quantifiers. In Woodford A. Beach, Samuel E. Fox & Shulamith Philosoph (eds.) *Papers from the 13th regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 372–383.

Miyagawa, Shigeru. 1997. *On the nature of WH-scope*. Manuscript, MIT.

Miyamoto, Yoichi & Yasuko Takata. 1998. Rigidity effects and the strong.weak wh features in SLA. In Annabel Greenhill, Mary Hughes, Heather Littlefield & Hugh Walsh (eds.) *Proceedings of the 22nd annual Boston University conference on language development, Volume 2*. Somerville MA: Cascadilla Press. 511–522.

Miyamoto, Yoichi & Maki Yamane. 1996. L2 rigidity: the scope principle in adult L2 grammar. Andy Stringfellow, Dalia Cahana-Amitay, Elizabeth Hughes & Andrea Zukowski (eds.) *Proceedings of the 20th annual Boston University conference on language development, Volume 2*. Somerville MA: Cascadilla Press. 494–505.

Montrul, Silvina, A. 1997. *Transitivity alternations in second language acquisition: A cross-linguistics study of English, Spanish and Turkish*. PhD dissertation: McGill University.

Nihongo Journal. 2002. Room sharing. *Nihongo Journal* 2002.6. 46–47.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer.

Nishigauchi, Taisuke. 1999. Quantification and wh-constructions. In Natsuko Tsujimura (ed.) *The handbook of Japanese linguistics*. Malden, MA & Oxford: Blackwell. 269–296.

Oller, John W. 1979. *Language tests at school: A pragmatic approach*. London: Longman.

Outcault, Samantha D., & Laurent Dekydtspotter. 2004. The resolution of scope ambiguity in English-French sentence interpretation. In Alejna Brugos, Linnea Micciulla & Christine E. Smith (eds.) *Proceedings of the 28th annual Boston University Conference on Language Development, Volume 2*. Somerville, MA: Cascadilla Press. 414–423.

Perlmutter, David. 1978. Impersonal passives and the unaccusative hypothesis. In Jeri J. Jaeger, Anthony C. Woodbury, Farrell Ackerman, Christine Chiarella, Orin D. Gensler, John Kingston, Eve F. Sweetser, Henry Thompson & Kenneth W. Whistler (eds.) *Proceedings of the fourth annual meeting of the Berkeley Linguistics Association*. Berkeley, CA: Berkeley Linguistics Association. 157–89.

Prince, Ellen. 1981. Toward a taxonomy of given/new information. In Peter Cole (ed.) *Radical pragmatics*. New York: Academic Press. 223–255.

Reinhart, Tanya. 1976. *The syntactic domain of anaphora*. PhD dissertation, MIT.

Reinhart, Tanya. 1978. Syntactic domains for semantic rules. In Franz Guenthner & S. J. Schmidt (eds.) *Formal semantics and pragmatics for natural languages*. Dordrecht: Reidel. 107–130.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation.* London: Croom Helm.

Reinhart, Tanya. 1995. Interface strategies. *OTS Working Papers* TL-95-002.

Reinhart, Tanya. 1997. Quantifier scope: how labor is divided between QR and choice functions. *Linguistics and Philosophy* 20: 335–397.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In Mark Baltin & Anthony S. Kroch (eds.) *Alternative conceptions of phrase structure*. Chicago: Chicago University Press. 182–200.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1: 69–118.

Sano, Tetsuya. 2004. Scope relations of QP's and scrambling in the acquisition of Japanese. In Jacqueline van Kampen & Sergio Baauw (eds.) *Proceedings of the GALA 2003 Conference on Language Acquisition*. 421–431.

Schwartz, Bonnie D. 1998. The second language instinct. *Lingua* 106: 133–160.

Schwartz, Bonnie D & Magda Gubala-Ryzak. 1992. Learnability and grammar reorganization in L2A: Against negative evidence causing the unlearning of verb movement. *Second Language Research* 8: 1–38.

Schwartz, Bonnie D. & Rex A. Sprouse. 1994. Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German Interlanguage. In Teun Hoekstra & Bonnie D. Schwartz (eds.) *Language acquisition studies in generative grammar: Papers in honor of Kenneth Wexler from the 1991 GLOW Workshops*. Amsterdam: John Benjamins. 317–368.

Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40–72.

Schwartz, Bonnie D. & Rex A. Sprouse. 2000. When syntactic theories evolve: Consequences for L2 acquisition research. In John Archibald (ed.) *Second language acquisition and linguistic theory*. Oxford: Blackwell. 156–186.

Schwarzschild, Roger 1999. GIVENness, Avoid F, and other constraints on the placement of accent. *Natural Language Semantics* 7: 141–177.

Shin, Kimie. 1983. Kuroozu-hou ni yoru nihongo nouryoku no sokutei (Using cloze tests to measure proficiency in Japanese). *Kyouiku kenkyuu* (Education research) 25. 141–178.

Shin, Kimie. 1987. Nihongo-nouryoku-hyouka no tame no hitokousatsu: Kurouzu tesuto no shinraisei, datousei, oyobi saitenhou no mondai (An examination of cloze tests in Japanese language proficiency evaluation: reliability, validity, and issues in scoring). *Gengo bunka ronshuu* (Papers on language and culture), *Volume 8.2*. Nagoya: Nagoya University Language Centre. 229–246.

Slabakova, Roumyana. 2001. *Telicity in the second language*. Amsterdam & Philadelphia, PA: John Benjamins.

Sorace, Antonella. 1996. The use of acceptability judgments in second language acquisition research. In William C. Ritchie & Tej K. Bhatia (eds.) *Handbook of second language acquisition*. San Diego: Academic Press. 375–409.

Sprouse, Rex. 2000. *What's inflection got to do with it?* Paper presented at the University of Durham, 16 October.

Suh, Jinhee. 1990. The scope interaction of wh-word and quantifier. In Hajime Hoji (ed.) *Japanese/Korean Linguistics*. Stanford, CA: CSLI. pp. 413–428.

Szabolcsi, Anna. 1994. Quantifiers in pair-list readings and the non-uniformity of quantification. In Paul Dekker & Martin Stokhof (eds.) *Proceedings of the Ninth Amsterdam Colloquium*. Amsterdam: ILLC/Department of Philosophy, University of Amsterdam. 645–664.

Szabolcsi, Anna. 1997. Strategies for scope taking. In Anna Szabolcsi (ed.) *Ways of scope taking*. Dordrecht: Kluwer. 109–155.

Szabolcsi, Anna. 2001. The syntax of scope. In Mark Baltin & Chris Collins (eds.) *Handbook of contemporary syntactic theory*. Oxford: Blackwell. 607–633.

Terada, Michiko. 1990. *Incorporation and argument structure in Japanese*. PhD dissertation, University of Massachusetts, Amherst.

Tohsaku, Yasu-hiko. 1994. *Yookoso! An invitation to contemporary Japanese*. New York: McGraw-Hill.

Tohsaku, Yasu-hiko. 1995. *Yookoso! Continuing with contemporary Japanese*. New York: McGraw-Hill.

Tomioka, Satoshi. 2004. *Pragmatics of LF intervention effects: Japanese and Korean wh-interrogatives*. Manuscript, University of Delaware.

Vallduví, Enric. 1992. *The informational component*. New York: Garland.

Vallduví, Enric. 1995. Structural properties of information packaging in Catalan. In Katalin É. Kiss (ed.) *Discourse configurational languages*. New York & Oxford: Oxford University Press. 122–153.

Vendler, Zeno. 1967. Each and every, any and all. Chapter 3 in *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press. 70–96.

White, Lydia. 1989a. The adjacency condition in Case assignment: Do L2 learners observe the Subset Principle? In Susan Gass & Jacqueline Schachter (eds.) *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press. 134–158.

White, Lydia. 1989b. *Universal Grammar and second language acquisition*. Amsterdam & Philadelphia, PA: John Benjamins.

White, Lydia. 1992. Subjacency violations and empty categories in L2 acquisition. In Helen Goodluck & Michael Rochemont (eds.) *Island constraints*. Dordrecht: Kluwer. 445–464.

Wible, David. 1990. *Subjects and the clausal structure of Chinese and English*. PhD dissertation, University of Illinois.

Yatsushiro, Kazuko. 1996. Case, scope and feature movement. In Masatoshi Koizumi, Masayuki Oishi & Uli Sauerland (eds.) *Formal approaches to Japanese linguistics 2: MIT Working Papers in Linguistics, Volume 29*. Cambridge, MA: Department of Linguistics, MIT. 319–335.

Yoshida, Keiko. 1993. *Syntax and semantics of wh-quantifier interactions*. PhD dissertation, Cornell University.

Yusa, Noriaki. 1995. Scope in Minimalist syntax. Review of Syntax of scope by Joseph Aoun and Yen-hui Audrey Li, The MIT Press, Cambridge, Massachusetts, 1993. *English Linguistics* 12: 319–346.