A Calculus for L1 Transfer*

William O’Grady

The phenomenon of transfer plays a prominent role in most approaches to SLA, including Focus on Form and the Interaction Hypothesis – to name two lines of inquiry that lay at the heart of Mike Long’s scholarship. The central thesis of this chapter is that transfer is best seen as a processing-driven strategy: L2 learners transfer operations from their L1 to the L2, unless those operations are more costly in the L2 than in the L1. Much of the chapter is devoted to the application of this idea to phenomena that arise when adult native speakers of English, Japanese, Korean, Spanish and Catalan go about acquiring a second or third language, producing transfer effects that have been difficult to accommodate in traditional approaches to cross-linguistic influence.

1. Introduction

Mike Long’s contributions to the study of second language acquisition are well documented and leave a legacy that will influence the field for decades to come. As Gass, Plonsky & Huntley (this volume) note in their bibliometric analysis of Mike’s influence, his interests encompassed several major lines of research – input, interaction, needs analysis, task-based language teaching, maturational constraints and ultimate attainment. I will focus here on transfer, a topic that interacts closely with those interests and is of broad relevance to the field of L2 research in general.

I will begin in the next section by outlining the approach to transfer that I have in mind and situating it within the larger explanatory framework that has come to be known as emergentism (e.g., MacWhinney & O’Grady 2015). In the subsequent five sections, I will review the results of several transfer-related studies which, taken together, point to a new perspective on this phenomenon. The chapter ends with some concluding remarks.

2. A Brief Overview

Well before the study of second language acquisition became a branch of linguistics, scholars were aware of transfer and its impact on the performance of adult L2 learners. An insightful observation along these lines was made more than a century ago by Harold Palmer, an early pioneer in the field:

[t]he adult will perceive dangerous analogies which will lead him astray. The child, whose reasoning faculties are comparatively underdeveloped, will not fall into these traps. (Palmer 1917:48)

Transfer continues to be a major topic of research in the field of second language acquisition. Indeed, it has become the primary concern of several contemporary approaches

---

* I am grateful for comments by Miho Choo, Kevin Gregg, an anonymous reviewer and the audience at the annual conference of the Korean Association of Teachers of English, as well as for many conversations in years past with Mike Long.
to SLA, including highly influential work within generative grammar (for succinct reviews, see Slabakova et al. 2020:13ff and Schwartz & Sprouse 2021a). I will begin by outlining the assumptions about language and learning that shape the approach to transfer adopted in this chapter.

2.1 Linguistic emergentism

The central tenet of linguistic emergentism is that language is a complex system in the technical sense, which means that its properties are shaped by the interaction of more basic forces. These forces appear to consist largely of processing pressures that seek to minimize the cost of the operations that are required for speech and comprehension (O’Grady 2005, 2015, 2021; Hawkins 2004, 2014). One such pressure involves a computational property that I will call Immediacy.¹

**Immediacy**

The processor favors operations that can be implemented without delay or the need for later revision.

The advantages of Immediacy are easy to see since operations that have this property avoid two potential obstacles to efficiency – delays that increase the burden on working memory and, in some cases, the need to revise a previously assigned interpretation. A simple example of an immediacy effect can be seen in the typology of *wh* questions in a language such as English.

Who did you meet _ yesterday?

Here, the empty position after the verb must be associated with the utterance-initial *wh* word in order for the sentence to be interpreted as a direct object question. Although this requires a certain amount of extra processing (e.g., Hofmeister & Sag 2010:380), its cost is minimized by the fact that the *wh* word is available as soon as the presence of a gap becomes evident.

![The gap can be interpreted immediately.](image)

Consistent with Immediacy, there is therefore no delay in interpreting the gap once it has been discovered.²

¹ This insight is widely recognized in the processing literature, where it goes by a number of different names including Maximize Online Processing (Hawkins 2004, 2015) and the Efficiency Requirement (O’Grady 2005), among others.

² This generalization is sometimes referred to as the Active Filler Hypothesis, so named because the ‘filler’ (the *wh* word) actively seeks a gap with which it can be associated (Clifton & Frazier 1989:292 & 297, Wagers & Phillips 2009:396-97, and many others).
Interestingly, there are no spoken languages in which *wh* words systematically occur to the right of the gap (Petronio & Lillo-Martin 1997:18, Hawkins 2014:30).

* read the book who?

This prohibition is arguably an *Immediacy* effect since patterns with sentence-final *wh* words force an obvious delay in the interpretation of the gap.

* read the book who?  
*↓?

*Immediate interpretation*  
of the gap is not possible.

I do not propose that *Immediacy* embodies an absolute requirement; in fact, it is well known that exceptions are possible (e.g., Hawkins 2014:215ff). The key claim is simply that instances of non-compliance add to processing cost, with consequences that can be discerned in phenomena ranging from typology to language acquisition, including the learning of a second language.

**2.2 Transfer and processing cost**

The leading idea that I will outline in this chapter is that transfer takes place in order to minimize processing cost. I’ll formulate this idea as follows, slightly revising the principle proposed by O’Grady (2010, 2012, 2013). (By ‘operation’ I mean any procedure involved in forming or interpreting a sentence: arranging words in a particular order, resolving a filler-gap dependency, finding an antecedent for a pronoun, and so on.)

**The Transfer Calculus**

L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

The Transfer Calculus embodies two simple claims.

i. Transfer takes place for a reason, which is to minimize processing cost by carrying familiar operations in the L1 over to the L2, where feasible.

ii. Transfer does not occur if it would have the effect of increasing processing cost in the L2.

On this view, the existence of transfer is accepted as a basic fact of second language acquisition, but it is not treated as either the default state of affairs or as an exception. Rather, its occurrence is determined by a fundamental calculus: transfer can occur if there is no resulting increase in processing cost, but not otherwise.

Before evaluating these claims, it is important to recognize that previous research has also recognized a connection between transfer and processing cost – a natural line of inquiry, as Filipović & Hawkins (2013) note.
Learners of a second language (L2) [should] prefer to minimize processing effort when using the grammatical and lexical properties of the L2, just as native speakers do.

An early and still very influential processing-based approach to transfer is the Competition Model pioneered by Bates & MacWhinney (1987). A key component of their theory involves ‘cue strength,’ which is shaped by the character of the input to which learners are exposed. In the case of English, word order is the most frequently used and most reliable clue for subjecthood; in Japanese case marking has priority; in Arabic animacy is a stronger cue than case marking (MacWhinney 2008:355); and so on. MacWhinney (2005:57) proposes that, by virtue of their entrenchment, cue rankings for the L1 are initially carried over to the L2, creating transfer effects such as the well-known oversensitivity to word order manifested by native speakers of English learning Japanese.

Another well-known processing-based approach to second language acquisition is the Developmentally Moderated Transfer Hypothesis put forward by Pienemann et al. (2005a,b) as part of what has come to be called Processability Theory.

Transfer is constrained by processability, in particular by the capacity of the L2 learner’s language processor... (Pienemann 2015:136)

On this view, transfer takes place in phrases before clauses, and in simple clauses before subordinate clauses, reflecting a sort of processing-cost hierarchy: phrase > clause > embedded clause.

Still another approach to transfer, more widely used in studies of L2 phonology than syntax, focuses on the possible role of typological markedness – a notion that is now often taken to be grounded in processing cost (Hawkins 2004, 2014; O’Grady 2021). Eckman (2010) summarizes the key insight underlying the markedness theory as follows:

… markedness in second language learning is a measure of relative difficulty and ease of transferability. The less marked a structure is, the easier it is to learn, and the less marked a construction is, the more likely it is to be transferred from the native to the target language. (Eckman 2010:490)

A classic example of syntactic markedness involves the NP Accessibility Hierarchy proposed by Keenan & Comrie (1977), which predicts (for example) that any language that allows indirect object relative clauses will also allow subject and direct object relative clauses.

Subject > Direct Object > Indirect Object > …

Subject RC: the student [that _ introduced Harry to Jane]
Direct object RC: the student [that Harry introduced _ to Jane]
Indirect object RC: the student [that Harry introduced Jane to _]
Doughty (1991) provided particularly striking evidence for the relevance of the Accessibility Hierarchy to second language acquisition. Working within a focus-on-form paradigm, she showed that instruction on a marked pattern (such as indirect object relative clauses) suffices to improve learners’ ability to produce less marked patterns (subject and direct object relative clauses). Later work extended these findings to the study of transfer effects (e.g. Hawkins 2007, Algady 2013, Xu 2014, Alroudhan & Ibn 2016, Alosaimi 2021).

In the rest of this chapter, I will pursue and extend the processing-based line of inquiry by examining the relevance of the Transfer Calculus to a series of phenomena that have thus far received relatively little attention in the literature. I will begin by considering a curious pattern involving verb deletion.

3. Verb Deletion in English and Japanese

A well-known phenomenon in the syntax of ellipsis involves the pattern illustrated below, in which the verb in the second conjunct of the English sentence can be dropped—an instance of phenomenon often called ‘gapping’ (e.g., Hwang 2020).

Gapping in English:

I **read** a book and [Sue **__** a magazine].

↑
verb  ↑
gap

The English pattern stands in sharp contrast to its counterpart in SOV languages such as Japanese (and Korean), in which the verb in the first conjunct is dropped.

Gapping in Japanese:

[Watashi-wa hon-o __] [Sue-wa magajin-o **yon-da**].

I-TOP book-ACC  Sue-TOP magazine-ACC read-PST

↑
gap

↑ verb

The two gapping patterns differ from each other in a fundamental way with regard to the principle of Immediacy, repeated here.

**Immediacy**

The processor favors operations that can be implemented without delay or the need for later revision.

Whereas the Verb–Gap pattern of English is compatible with this preference, its Gap–Verb counterpart in Japanese is not.
Verb–Gap (English)
I read a book and [Sue ___ a magazine].

\[ \text{The gap can be interpreted immediately.} \]

Gap–Verb (Japanese)
\[ \text{I-top book-acc Sue-top magazine-acc read-pst} \]

\[ \text{Immediate interpretation of the gap is not possible.} \]

These differences are highly relevant to the metric on which the Transfer Calculus is based.

**The Transfer Calculus**

L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

The Transfer Calculus makes two predictions about gapping, one for English speakers learning Japanese and the other for Japanese speakers learning English. Let us consider each in turn.

### 3.1 English-speaking learners of Japanese

The prediction that the Transfer Calculus makes for English-speaking learners of Japanese can be formulated as follows:

**Prediction 1 (Verb deletion in Japanese)**

English-speaking learners should prefer the Verb–Gap pattern of their L1 over the Gap–Verb pattern in Japanese.

The basis for this prediction is not simply that English happens to make use of the Verb–Gap pattern and that its properties are automatically carried over to the second language. It is that the Verb–Gap pattern incurs a lower cost than its Japanese counterpart since the missing verb can be immediately identified at the point where the gap is encountered by simple reference to the first conjunct.

The Verb–Gap pattern:
\[ \text{NP V NP and NP ___ NP} \]

\[ \text{The gap can be interpreted immediately.} \]

By transferring this pattern to Japanese, learners reduce the cost of gapping in their second language, as depicted below.
After transfer to Japanese:

\[
\begin{array}{c}
\text{NP NP V} \quad \text{NP NP } \quad \uparrow \\
\text{The gap can be interpreted immediately.}
\end{array}
\]

Cost-wise, this is an obvious improvement over the Japanese pattern, in which interpretation of the missing verb has to be delayed, in violation of Immediacy.

\[
\begin{array}{c}
\text{NP NP } \quad \text{NP NP V} \\
\downarrow \\
\text{Immediate interpretation of the gap is not possible.}
\end{array}
\]

Data from an experiment conducted by O’Grady (1999) is relevant to the assessment of this scenario.

**Participants:**
75 English-speaking college students studying Japanese as an L2
(20 second semester, 32 third semester, 23 fourth semester)

**Method and materials:**
- Subjects were asked to judge the acceptability of gapping patterns by rating them on the following scale.

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & \text{I don’t know}
\end{array}
\]

- There were 5 tokens of the Verb–Gap pattern, and 5 tokens of the Gap–Verb pattern.\(^3\)

**Sample Gap–Verb pattern (acceptable in Japanese):**
I was asked what Suzuki drank and what Tanaka drank. Here is what I think.

\[
\begin{array}{c}
\text{Suzuki-} \text{HON-TOP beer-ACC } \text{Tanaka-} \text{HON-TOP soda-ACC drink-PST that think-PRS}
\end{array}
\]

**Sample Verb -Gap pattern (unacceptable in Japanese):**
I was asked what Suzuki drank and what Tanaka drank. Here is what I think.

\[
\begin{array}{c}
*\text{Suzuki-} \text{HON-TOP beer-ACC drink-CON Tanaka-} \text{HON-TOP soda-ACC that think-PRS}
\end{array}
\]

**Results:**
As predicted, English-speaking learners show a preference for the low-cost Verb–Gap pattern even though it is unacceptable in Japanese. As reported below, that pattern received

\(^3\) HON = honorific, TOP = topic, ACC = accusative, PST = past, CON = connective. Because the JSL learners were not yet fully fluent in Japanese, the context was presented in English; however, the test sentences were presented in Japanese in the standard kana-kanji script.
a mean rating of 3.09 (out of 5) compared to just 2.36 for the Gap–Verb pattern. The difference is statistically significant (p = .000).

<table>
<thead>
<tr>
<th>Gap-Verb pattern (✓ in Japanese)</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Watashi-wa hon-o __ ] [Sue-wa magajin-o yon-da].</td>
<td>2.36</td>
</tr>
<tr>
<td>I-TOP book-ACC __ Sue-TOP magazine-ACC read-PST</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb-Gap pattern (*) in Japanese</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[Watashi-wa hon-o yom-i ] [Sue-wa magajin-o __].</td>
<td>3.09</td>
</tr>
<tr>
<td>I-TOP book-ACC read-CON Sue-TOP magazine-ACC __</td>
<td></td>
</tr>
</tbody>
</table>

Now let us consider the reverse situation involving the acquisition of English by Japanese speakers, for which the Transfer Calculus makes a quite different prediction.

3.2 Japanese-speaking learners of English

If the Transfer Calculus is right, Japanese speakers should be reluctant to transfer the Gap–V pattern of their first language to English because of its extra cost compared to the V–Gap option offered by the L2.

The Gap-Verb pattern of Japanese transferred to English:

\[
\begin{array}{c}
\text{[NP __ NP] and [NP \ V \ NP]} \\
\text{Immediate interpretation} \\
\text{of the gap is not possible.}
\end{array}
\]

The Verb-Gap pattern in English:

\[
\begin{array}{c}
\text{[NP \ V \ NP] and [NP \ __ \ NP]} \\
\text{The gap can be interpreted immediately.}
\end{array}
\]

The Transfer Calculus therefore makes the following prediction.

**Prediction 2 (Verb deletion in English)**

Japanese-speaking learners should prefer the Verb–Gap pattern of English over the Gap–Verb pattern of their first language.

Data from the afore-mentioned 1999 study sheds light on this issue as well.
Participants:
34 Japanese native speakers learning English as an L2
(22 first-year students, 12 third-year students)

Method and materials:
• A written acceptability judgment task similar to the one in the first experiment was employed.4
• There were 5 tokens of the Verb–Gap pattern, and 5 tokens of the Gap–Verb pattern, as exemplified below.

  Sample V–Gap pattern (acceptable in English):
  Peter asked me what Susan ate and what Harvey ate.
  I said that [Susan ate a hamburger] and [Harvey __ a sandwich].

  Sample Gap–V pattern (unacceptable in English):
  Peter asked me what Susan ate and what Harvey ate.
  *I said that [Susan __ a hamburger] and [Harvey ate a sandwich].

Results:
As predicted, Japanese-speaking L2 learners show a preference for the low-cost Verb–Gap pattern of English (mean rating 2.33) over the Gap–Verb pattern of their first language (1.75). The difference is statistically significant (p .0022).

<table>
<thead>
<tr>
<th>Verb-Gap pattern (√ in English)</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>[I read a book] and [Sue __ a magazine].</td>
<td>2.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gap–verb pattern (* in English)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[I __ a book] and [Sue read a magazine].</td>
</tr>
</tbody>
</table>

3.3 Implications
To review, there is an evident asymmetry in our results:
• Native speakers of English prefer the Verb–Gap pattern of their first language when assessing gapping in Japanese.
• Native speakers of Japanese are reluctant to accept the Gap–Verb pattern of their first language when assessing gapping in English.
This asymmetry makes perfect sense from a cost-based perspective. The preference by English speakers for the V–Gap pattern in Japanese facilitates the processing of ellipsis in that language, just as it does in English.

---

4 Because the Japanese learners were not yet fully fluent in English, the context was presented in Japanese.
The V–Gap pattern transferred to Japanese:

\[
\text{[NP } \text{NP } V \text{]} \ldots \text{[NP } \text{NP } \_	ext{]} \quad \uparrow
\]

The gap can be interpreted immediately.

In contrast, adopting the Gap–Verb pattern of Japanese in English has the opposite effect, precluding immediate interpretation of the gap.

The Gap–Verb pattern transferred to English:

\[
\text{[NP } \_	ext{ NP]} \ldots \text{[NP } V \text{ NP]} \quad \downarrow
\]

Immediate interpretation
of the gap is not possible.

In sum, the predictions of the Transfer Calculus appear to be borne out in the case of gapping: the likelihood of transfer is driven by a sensitivity to processing cost. A similar effect can be discerned in the very different phenomenon to which we turn next.

4. The Interpretation of Quantified NPs in English and Korean

A striking difference between English and Korean involves the preferred interpretation of sentences in which negation interacts with a universal quantifier in the direct object position, creating the relationship known as ‘scope.’

English pattern:
The boy didn’t catch all the fish. (= ‘The boy caught only some of the fish.’)

Korean pattern:
Sonyen-i motun punge-lul an cap-ass-ta.
boy-NOM all fish-ACC not catch-PST-DECL
‘The boy didn’t catch all the fish.’ (= ‘All the fish went uncaught.’)

The two languages differ with respect to the interpretation of these patterns, as described in detail by O’Grady (2013). Once again, the key variable involves Immediacy.

Immediacy
The processor favors operations that can be implemented without delay or the need for later revision.

The English preference for the negated (‘not all’) interpretation of the QNP, in which the boy caught just some fish, is fully consistent with Immediacy. At the point at which the processor comes upon the quantified NP (QNP) all the fish, the previously encountered negative (not) is available – allowing a negated interpretation to be derived without delay.
The negated interpretation of the QNP (English):
The boy didn’t catch all the fish.

↑
Negated interpretation of the QNP is immediately available.

CAUGHT NOT CAUGHT

In contrast, we find a very different situation in Korean.

The non-negated interpretation of the QNP (Korean):

Sonyen-i motun punge-lul an cap-ass-ta.
boy-NOM all fish-ACC not catch-PST-DECL

↑
The QNP is interpreted before the negative is encountered.

As illustrated here, the QNP motun punge-lul ‘all the fish’ is encountered before it can fall under the influence of the negative. In contrast to its English counterpart, the sentence should therefore describe a situation in which all the fish went uncaught. Several experimental studies have confirmed that this interpretation is in fact strongly preferred by native speakers of Korean (Han et al. 2007, Lee 2009, O’Grady et al. 2009, 2011).

The difference between the English and Korean patterns lays the groundwork for a further test of the Transfer Calculus, repeated here.

The Transfer Calculus
L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

Two predictions can be made – one for English-speaking learners of Korean and the other for Korean learners of English. Let us consider each in turn.

4.1 English-speaking learners of Korean

A first prediction involves the interpretation of QNPs in Korean by native speakers of English.

Prediction 3 (QNPs in Korean)
English-speaking learners should prefer the non-negated interpretation of the QNP that is also preferred by native speakers of Korean. They should reject the negated interpretation that is dominant in their first language, English.

The basis for this prediction lies in Immediacy. In contrast to English (see the preceding page), the QNP occurs prior to the negative in Korean
Interpretation of the QNP in Korean:

The boy **all the fish** not caught.

↑

*The QNP is interpreted before the negative is encountered; there is no immediate opportunity to derive the negated interpretation.*

A negated ‘not all’ reading in this case would call for reinterpretation of the QNP, contra the principle of Immediacy. This in turn would require a more costly interpretive procedure than the one employed in the L1 (English). Transfer is thus ruled out. Put simply, if the Transfer Calculus is right, English speakers are better off interpreting the Korean sentence as if they were Koreans.

The prediction is confirmed by the results of an experiment conducted by O’Grady (2013).

**Participants:**
10 native speakers of English in their late teens or early twenties who were enrolled in a 4th-semester course in Korean as a second language at an American university. (None were heritage learners of Korean.)

**Method and materials:**
Participants were asked to match a test sentence with an appropriate contextual passage, one of which was compatible with a negated interpretation of the QNP and the other with a non-negated interpretation. There were ten test items in all. Here is a sample.

| Tom-i motun chayk-ul an ilk-ess-ta. | Context 1 (negated interpretation):  
Tom-NOM all book-ACC not read  
I gave Tom all the books that he was supposed to read, but he read only some of them. | Context 2 (non-negated interpretation):  
I gave Tom all the books that he was supposed to read, but he didn’t read any of them. |

Participants who prefer the negated interpretation of the QNP should match the sample test sentence with the context on the left, in which not all the books were read. In contrast, those who favor a non-negated interpretation should select the context on the right, in which all the books go unread.

**Results:**
Native speakers of English were strongly adverse to the negated interpretation of the QNP in Korean, favoring the less costly non-negated reading 100% of the time. As predicted, there are no signs of transfer of any sort.

**4.2 Korean-speaking learners of English**
If the theory of cost-based transfer is right, Korean-speaking learners of L2 English should prefer the non-negated interpretation of the QNP that is also favored in Korean.
Prediction 4 (QNPs in English)
Korean-speaking learners should prefer the non-negated interpretation of the QNP that is dominant in their L1 over the negated interpretation that is dominant in English.

This prediction follows from the central tenet of the Transfer Calculus, which is that learners should carry the procedures from their L1 over to the L2 as long as there is no increase in processing cost.

The Transfer Calculus
L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

Crucially, the procedure that yields a non-negated interpretation for the QNP in Korean can be applied to English with no increase in cost.

How the QNP is interpreted by Koreans in their L1:
The boy all the fish not caught.

The QNP is interpreted here with no regard for the negative.

How the QNP is interpreted by Koreans in L2 English:
The boy didn’t catch all the fish.

The QNP is interpreted here with no regard for the negative.

Even though not precedes the QNP in English and is therefore available to derive the negated interpretation, nothing requires that this actually happen. Indeed, children learning English as a first language commonly ignore the negative when interpreting the QNP (Musolino et al. 2000, Musolino & Lidz 2006). The Transfer Calculus predicts that Korean-speaking learners of English will adopt a similar course of action, at least initially. Put simply, they should interpret English as if it were Korean.

Key evidence comes from an experiment conducted by O’Grady, Kwak, Lee & Lee (2011).

Participants:
42 native speakers of Korean at a university in Seoul (intermediate to high-intermediate learners of English)

Method and materials:
• Truth Value Judgment task
• 4 test items in which a negated interpretation of the QNP was favored and 4 in which a non-negated reading should be preferred.
The girl didn’t throw out all the boxes.

Figure 1. A sample test item

If the L2 learners adopt the negated ‘not all’ interpretation of the QNP, they should accept the sentence in Figure 1 as a true description of the depicted situation (since not all of the boxes were thrown out). On the other hand, if they adopt the non-negated interpretation (according to which all boxes remain inside), they should reject the sentence as false.

Results:
The Korean learners of English showed a strong preference for the non-negated reading of the QNP, rejecting sentences such as the one above as false 72% of the time. When tested on similar items in their native Korean, they rejected the sentence as false 79% of the time. The similarity of the scores in the two conditions strongly suggests that the same interpretive mechanism is at work – consistent with the predicted transfer effect.5

4.3 Implications

In sum, as was the case with verb deletion, we see a major and intriguing asymmetry in the workings of transfer in the sentences that involve the interaction of a negative and a universal quantifier.

• Native English speakers avoid transfer when learning Korean: as we have seen, they favor the non-negated reading of the QNP rather than the ‘not all’ reading typical of English. As explained in Section 4.1, this happens because the ‘not all’ interpretation cannot be derived in Korean without violating Immediacy, making it more costly in the L2 than in the L1.

Interpretation of the QNP in Korean:
The boy all the fish not caught.

The QNP is interpreted before the negative is encountered; there is no immediate opportunity to derive the negated interpretation.

5 Native speakers of English accept such sentences as true 100% of the time (Musolino et al. 2000, Musolino & Lidz 2006).
In contrast, native Korean speakers carry the non-negated interpretation preferred in their native language when interpreting English sentences—a result that suggests transfer.

How can this be? The answer is in fact quite straightforward. Both developmental trajectories are perfectly aligned with the metric underlying the Transfer Calculus, reflecting the simple fact that learners seek to minimize processing cost in their L2.

As we will see next, a different scopal phenomenon—this time in Japanese—offers still further support for a theory of cost-related transfer.

5. The Interpretation of Indefinite NPs in English and Japanese

The two previous sections of this chapter have dealt with what might be called ‘bi-directional transfer’ since they focus on the comparison of mirror-image learning scenarios: the acquisition of language x by speakers of language y and the acquisition of language y by speakers of language x. In this section, I examine the more usual context for the study of transfer, which involves the acquisition of a single particular second language by native speakers of another particular language.

5.1 The phenomenon

The particular phenomenon on which I will focus involves the interpretation of indefinite pronouns in patterns that also contain a universal quantifier. The following English sentence is a case in point.

Someone stroked every cat.

\[
\begin{align*}
\text{Someone} & \quad \uparrow \\
\text{ indefinite pronoun} & \quad \text{universal quantifier} \\
\text{every cat} & \quad \uparrow
\end{align*}
\]

The dominant interpretation of this sentence is that a particular person stroked a number of cats. This reading is derived by interpreting each NP as it is encountered, consistent with Immediacy.

Immediacy

The processor favors operations that can be implemented without delay or the need for later revision.

According to this scenario, the processor first encounters someone, which it takes to pick out a single person; later, it comes upon every cat, which it takes to denote multiple felines. Each NP is interpreted as soon as it is encountered, and neither influences the interpretation of the other. I will call this the ‘specific interpretation’ since someone is taken to refer to a single particular individual.
The specific interpretation of an indefinite NP:

Someone stroked every cat.

A second and less common interpretation is also possible: different people are involved in stroking the cats. As depicted below, this reading involves reinterpretation of the indefinite NP upon encountering the QNP *every cat*, so that there is (for instance) a different girl for each cat. This is often called the ‘distributed interpretation.’

The distributed interpretation of an indefinite NP:

Someone stroked every cat.

By the standards of Immediacy, the distributed interpretation should be more costly since it requires reinterpretation of the subject NP. This seems to be correct: as the literature on quantifiers notes, even native speakers find it difficult to derive the distributed reading.

[the distributed interpretation] is a marked option: It is often very hard to obtain and it requires a strong discourse motivation. (Reinhart 1997:370)

Moreover, there is evidence that (all other things being equal) the distributed interpretation takes longer to process (Anderson 2004) and is actually prohibited in many languages (Keenan 1974:301-02, 1976:319; Hawkins 2004:17).

Japanese is one of the many languages that do not permit a distributed interpretation in the patterns under consideration.

Dareka-ga dono neko-mo nade-ta.
someone-NOM every cat-FOC stroke-PST
‘Someone stroked every cat.’
(sole interpretation = ‘One specific person stroked all the cats.’)

The contrast between English and Japanese raises an obvious question: how do native speakers of English go about interpreting Japanese sentences like the one above?

---

6 Drawing on a different terminology, Marsden uses the term ‘forward scope’ for the specific interpretation and ‘inverse scope’ for the distributed interpretation.
5.2 The interpretation of *dareka* by native speakers of English

On the view I propose, the answer for the question at hand should follow from the Transfer Calculus, repeated here.

**The Transfer Calculus**

L2 learners transfer operations from their L1 to the L2 – unless those operations are more costly in the L2 than in the L1.

The specific interpretation of an indefinite NP should therefore be relatively easy in Japanese, just as it is in English, since Immediacy is satisfied in this case in both languages. As illustrated below, the specific interpretation in Japanese involves picking out a single girl upon encountering *dareka*-ga ‘someone’ and a set of cats upon encountering *dono neko-mo* ‘every cat’ – with neither NP affecting the interpretation of the other.

The specific interpretation of an indefinite NP in Japanese:

(compatible with Immediacy, as in English):

Dareka-ga  dono neko-mo  nade-ta.
someone-NOM  every  cat-FOC  stroke-PST

↑↑

‘Someone stroked every cat.’ (= ‘One specific person stroked all the cats.’)

A distributed interpretation – were it allowed in Japanese – would be more costly than the specific interpretation, but no more costly than the distributed reading in English. In fact, the relevant operations would be essentially identical to those used in the earlier English example.

The distributed interpretation of an indefinite NP in Japanese:

(incompatible with Immediacy, as in English):

Dareka-ga  dono neko-mo  nade-ta.
someone-NOM  every  cat-FOC  stroke-PST

↑↑

*reinterpretation of dareka*

Given these facts, the Transfer Calculus predicts that the procedures required for both the specific interpretation and the distributed interpretation should be transferable from English to Japanese.
**Prediction 5 (Indefinite NPs in Japanese)**

English-speaking learners of Japanese should permit both the specific and distributed interpretation of indefinite NPs in Japanese.

Marsden (2004) conducted an experiment to investigate how native speakers of English interpret *dareka* in the sort of sentences that we are considering.

**Participants:**
- 18 intermediate English-speaking learners of Japanese as a second language
- 9 advanced English-speaking learners of Japanese as a second language

**Method and materials:**
Participants were asked to indicate how well a particular sentence matched one or the other of the pictures below by ranking it on a scale of 0 to 3. (A score of 1.5 or higher was taken to indicate acceptance of the relevant interpretation.)

![Sample pictures from Marsden's study](p. 123)

**Results:**
The table below summarizes Marsden’s results (p. 195) for English-speaking learners of Japanese.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Specific</th>
<th>Distributed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>2.74</td>
<td>1.61</td>
</tr>
<tr>
<td>Advanced</td>
<td>2.82</td>
<td>1.38</td>
</tr>
<tr>
<td>Native speakers of Japanese</td>
<td>2.26</td>
<td>0.69</td>
</tr>
</tbody>
</table>

**5.3 Implications**

Two findings support the prediction of the Transfer Calculus.

i. The L2 learners assign very high ratings to the specific interpretation of *dareka*. Since that reading complies with Immediacy and therefore has a low cost in both languages, the underlying interpretive operation should be carried over from English to Japanese. This seems to be the case.
ii. The L2 learners assign moderately high ratings to the distributed interpretation of *dareka*, especially compared to rating of 0.69 that it received from native speakers of Japanese. Although this interpretation is costly in English, there is (as noted previously) no reason to think that it would be inherently more costly in Japanese if it were allowed there. Transfer should therefore take place, as it indeed does.

There is nonetheless a potential mystery here.

The ratings that native speakers of English assign to the specific interpretation in Japanese are very close to the one that they assign to the comparable interpretation in English (2.68); see table 2.

<table>
<thead>
<tr>
<th>Language and Participants</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate L2 learners judging Japanese</td>
<td>2.74</td>
</tr>
<tr>
<td>Advanced L2 learners judging Japanese</td>
<td>2.82</td>
</tr>
<tr>
<td>Native speakers of English judging English</td>
<td>2.68</td>
</tr>
</tbody>
</table>

In contrast, the ratings that they assign to the distributed interpretation in Japanese (1.61 and 1.38) are substantially lower than the one that they give for that interpretation in English (1.96); see table 3.

<table>
<thead>
<tr>
<th>Language and Participants</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate L2 learners judging Japanese</td>
<td>1.61</td>
</tr>
<tr>
<td>Advanced L2 learners judging Japanese</td>
<td>1.38</td>
</tr>
<tr>
<td>Native speakers of English judging English</td>
<td>1.96</td>
</tr>
</tbody>
</table>

Why is this the case?

5.4 The Weakness Corollary

As I see it, the puzzle is best subsumed under a more general question, which can be formulated as follows.

Why are costly operations more difficult to transfer than their less costly counterparts?

The answer, I believe, lies in a supplementary principle that, like the Transfer Calculus, is motivated by processing pressures.\(^7\)

---

\(^7\) I use the term ‘corollary’ here in the non-technical sense of the Merriam-Webster dictionary: ‘something that naturally follows; something that incidentally or naturally accompanies’ another proposition.
The Weakness Corollary
Weak L1 routines become even more feeble when carried over to an L2.

This idea fits well with the widely accepted finding that processing in a second language is slower and less efficient than in one’s first language.

… what do studies of the aging brain reveal about L2 acquisition and processing? From the cognitive literature, we learn that the associative memory and incremental learning elements of language learning are steadily compromised by age, as are the working memory and processing speed components of language processing and production. (Birdsong 2006:34)

For all three factors (working memory capacity, decoding ability, and speed of processing) there is evidence that late L2 learners have poorer abilities than native speakers. (McDonald 2006:385)

If this is right, then the transferred version of the procedure that English-speaking learners use for the distributed interpretation should be doubly weak. It is costly in the learners’ first language, and the computational resources required to implement it in the second language are compromised.

Taken together, these considerations lead to the following prediction by the Weakness Corollary.

PREDICTION 6 (DISTRIBUTED INTERPRETATION OF INDEFINITE NPS IN JAPANESE)
English-speaking learners should rate the distributed interpretation of Japanese dareka less favorably than the distributed interpretation of English someone.

Marsden’s results appear to confirm this prediction, as the data in table 1 – 3 show: there is a strong tendency to accept the specific interpretation of the indefinite pronoun and a weaker tendency to permit the distributed reading.8

In sum, the developmental facts that surround the interpretation of indefinite NPs in Japanese fit well with the Transfer Calculus.

• English-speaking learners welcome the specific interpretation of dareka in Japanese, since that reading involves the same low-cost interpretive procedure that is used in English; there is no additional cost when it is employed in Japanese.

• English-speaking learners show low-to-marginal rates of acceptance for the distributed reading in Japanese, for which their first language provides a costly and

---

8 A third effect may also be in play. Because L2 learners never encounter the distributed interpretation in Japanese (where it is banned), the underlying operation will never be activated – raising the possibility that it might wane to the point of inaccessibility. Marsden’s data suggests that something like this may in fact happen. Whereas learners with intermediate proficiency in Japanese tended to permit the distributed interpretation at least some of the time, six of the nine speakers with advanced proficiency ‘consistently rejected’ it (p. 208). Significantly, members of that latter group had spent substantially more time in Japan, thereby increasing the period during which the distributed derivation went unactivated.
little-used interpretive operation. Although that procedure is transferred to Japanese, it is less accessible in the second language, consistent with the Weakness Corollary.

As we will see next, the Weakness Corollary has a role to play in a quite different phenomenon – this one involving the acquisition of Mandarin by native speakers of Japanese.

6. The Interpretation of Null Arguments in Japanese and Chinese

A well-documented feature of many languages is the possibility of null arguments, especially in subject position. Although unpronounced subjects were once thought to be essentially alike in all languages that allowed them, later work uncovered striking differences that open the door to intriguing lines of inquiry in the field of second language research. One such investigation is particularly relevant to the view of transfer that I have been developing.

6.1 The phenomenon

Monou (2013) reports on a little-known fact that arises in contexts such as the following in Japanese.

San-nin-no keisatsu-ga Sato-san-no ie-ni kita.
three-CL-GEN police-SUBJ Sato-HON-GEN house-to came
‘Three police officers came to Ms Sato’s house.’

pro Yamada-san-no ie-ni-mo kita.
Yamada-HON-GEN house-to-also came
‘[They] also came to Ms Yamada’s house.’

On one interpretation, the null subject pronoun (represented as pro) in the second sentence refers to the set of police officers to which reference is made in the first sentence. I will refer to this as the ‘definite’ interpretation, which is illustrated in the figure below by depicting the same set of police officers twice – once at Ms Sato’s house and once at Ms Yamada’s house.

![Figure 3. The definite interpretation as depicted by Monou](image)

A second reading, which I will call the ‘indefinite’ interpretation, is more demanding since it requires the construction of a new set of police officers, as depicted in Figure 4.
In contrast to the definite interpretation, this second reading does not satisfy Immediacy since the reference of the null pronoun cannot be determined by simply calling up the set of police officers alluded to in the first sentence of the test item. An entirely new set of officers must be conjured.

The definite interpretation (the referents of pro are immediately available):

\[
\begin{align*}
\text{pro } & \text{Yamada-san-no ie-ni-mo kita.} \\
\text{immediate interpretation} & \downarrow \\
& \text{by reference to the previously mentioned set of police officers}
\end{align*}
\]

The indefinite interpretation (the referents of pro are not immediately available):

\[
\begin{align*}
\text{pro } & \text{Yamada-san-no ie-ni-mo kita.} \\
? & \downarrow \\
& \text{Immediate interpretation} \\
& \text{is not possible; a new set of police officers must be constructed.}
\end{align*}
\]

We can therefore conclude that the indefinite interpretation is more costly – an assessment for which there is also typological evidence. As Monou notes (p. 96), some null-subject languages allow only the definite interpretation, but no language allows only the indefinite interpretation.

One language that allows just the definite interpretation is Mandarin, as illustrated in the following example from Monou (p. 9).

Sān-gè jīngchá lái-le Zuōténg jiā. three-CL police officer come-ASP Ms Sato’s house
‘Three police officers came to Ms Sato’s house.’

\[
\begin{align*}
\text{pro } & \text{yē lái-le Shāntián jiā.} \\
\text{also come-ASP Ms Yamada’s house} & \downarrow \\
& \text{They also came to Ms Yamada’s house. [definite interpretation only]}
\end{align*}
\]
A question now arises: what can be expected to happen when Japanese speakers learn Mandarin as a second language?

6.2 The interpretation of Chinese pro by native speakers of Japanese

Given the Transfer Calculus and the Weakness Corollary, two predictions can be made about how native speakers of Japanese will go about interpreting pro in Mandarin.

**Prediction 7 (Definite Interpretation of pro in Mandarin)**
Consistent with the Transfer Calculus, Japanese-speaking learners should accept the definite interpretation of pro in Mandarin since the relevant interpretive procedure can be carried over from Japanese at no additional cost.

**Prediction 8 (Indefinite Interpretation of pro in Mandarin)**
Consistent with the Weakness Corollary, Japanese-speaking learners should rate the more costly indefinite interpretation of pro in Mandarin substantially lower than in their first language.

We are able to test these predictions with the help of data from an experiment conducted by Monou for a different purpose.

*Participants:*
22 undergraduate students at a university in Tokyo, all of whom were native speakers of Japanese. On average, they had been studying Mandarin at their university for 11 months. A pre-test revealed that all 22 participants accepted both the definite and indefinite interpretations of pro in their first language.

*Method and materials:*
  * Truth Value Judgment Task
  * 5 sets of test items such as the one illustrated above, each of which was matched with pictures to test its compatibility with a definite interpretation and an indefinite interpretation.

*Results:*
Monou (p. 36) reports a sharp contrast in the rate at which the two interpretations were accepted by the Japanese-speaking learners of Mandarin.

<table>
<thead>
<tr>
<th>Definite Interpretation (same 3 police officers)</th>
<th>Indefinite Interpretation (different set of 3 police officers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>86.4%</td>
<td>35.5%</td>
</tr>
</tbody>
</table>
Not only is the acceptance rate much lower for the indefinite interpretation, 10 of Monou’s 22 participants rejected the truth of all the Mandarin test items that occurred in contexts calling for the indefinite reading (p. 41).

6.3 Implications

Monou’s results are expected if the Transfer Calculus interacts with the Weakness Corollary in the way that I have proposed, creating the following developmental scenario.

- Consistent with the Transfer Calculus, the low-cost and frequently instantiated procedure associated with the definite interpretation is carried over to Mandarin since it is no more costly in that language.

- Consistent with the Weakness Corollary, the high-cost and infrequently activated procedure associated with the indefinite interpretation is transferred to the second language in a very weak form, where it is highly susceptible to demise.

This is exactly what Monou’s findings appear to show.

7. A Note on Third-Language Acquisition

In recent years, the study of transfer has moved in a new and challenging direction to include the acquisition of third languages (e.g., Rothman, González & Puig-Mayenco 2019, Puig-Mayenco, González & Rothman 2020, Schwartz & Sprouse 2021a,b). The key question in this endeavor turns on which of two earlier acquired languages has the greater influence on the L3 – a matter to which I will now turn.

7.1 The Phenomenon

I will focus here on an intriguing contrast involving negative concord – a phenomenon that requires a negative pronoun to be accompanied by a sentential negative, as in the following example from Catalan.9

Pre-verbal negative pronoun with a sentential negative in Catalan:

Ningú no beu cafè.
nobody not drink coffee
‘Nobody drinks coffee.’

Despite the presence of two negatives, the sentence has a ‘single-negative’ interpretation: nobody drinks coffee.

In contrast, the comparable pattern in Spanish has a ‘double-negative’ interpretation that essentially cancels negation.

---

9 In fact, the situation appears to be somewhat more complex. Based on experimental data, Déprez et al. (2015) report that one ‘population’ of Catalan speakers treats the sentential negative in this pattern as optional.
Pre-verbal negative pronoun with a sentential negative in Spanish:

Nadie no bebe café.

nobody not drink coffee

‘There is nobody that does not drink coffee.’ (= ‘Everyone drinks coffee.’)

A question that now arises has to do with how bilingual speakers of Catalan and Spanish will behave when learning English as a third language. In particular, which pattern of negation will they transfer to English – the one from Catalan or the one from Spanish? Rothman (2015) suggests that the key variable is the ‘structural proximity’ of the three languages. According to his Typological Proximity Model, transfer comes from the previously learned language that is typologically most similar to the L3 (see, e.g., Rothman et al. 2019:154ff).

Puig-Mayenco & Rothman (2020) argue that Catalan is typologically more like English than is Spanish, based on certain phonological and phonotactic similarities identified by Puig-Mayenco & Marsden (2018:508). This leads to the following prediction:

PREDICTION OF THE TYPOLOGICAL PROXIMITY MODEL
Catalan-Spanish bilinguals should transfer features of Catalan rather than Spanish to English, regardless of which language they learned first.

7.2 The acquisition of English negation by Catalan and Spanish bilinguals
Puig-Mayenco & Rothman tested their prediction with the help of an experiment involving two groups of Catalan-Spanish bilinguals who had just completed a two-month English-language course specifically designed for beginners.

Participants:
- 35 L1-Catalan, L2-Spanish bilinguals (mean age = 54.4)
- 24 L1-Spanish, L2-Catalan bilinguals (mean age = 52.1)

Method and materials:
In order to determine which of the two languages has more influence on the acquisition of English as a third language, Puig-Mayenco & Rothman tested the two groups of learners on unfamiliar sentences such as the following.

Nobody doesn’t drink coffee.

The task involved choosing between a picture in which no one is drinking coffee and a picture in which there is no one who is not drinking coffee.

The English sentence permits only a double-negative interpretation (‘There’s nobody that doesn’t drink coffee’), but this is irrelevant. What matters for the purposes of the experiment is what the learners think that the sentence means.

Results:
Table 5 presents Puig-Mayenco & Rothman’s results.
Table 5. Preferred interpretation of the English sentences

<table>
<thead>
<tr>
<th>Participants</th>
<th>Single-Negative Interpretation</th>
<th>Double-Negative Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-Catalan, L2-Spanish</td>
<td>90%</td>
<td>10%</td>
</tr>
<tr>
<td>L1-Spanish, L2-Catalan</td>
<td>78%</td>
<td>22%</td>
</tr>
</tbody>
</table>

a ‘Nobody drinks coffee.’
b ‘There’s nobody that doesn’t drink coffee.’

Crucially, both the L1-Spanish speakers and the L1-Catalan speakers prefer the single-negative interpretation (‘Nobody drinks coffee’) – the reading that is associated with the corresponding Catalan sentence rather than its Spanish counterpart.

7.3 Implications

Puig-Mayenco & Rothman interpret their result as evidence for transfer from Catalan, in support of their Typological Proximity Model. But their conclusion is not without problems, since a precise metric for typological proximity remains to be developed. Interestingly, Immediacy offers an alternative explanation for why the single-negative interpretation is preferred.

**Immediacy**

The processor favors operations that can be implemented without delay or the need for later revision.

The key observation is that only the single-negative reading of the English sentence allows an immediate and stable interpretation of the negative pronoun *nobody*.\(^{10}\) (I use the symbol Ø to indicate a null set.)

```
Nobody Ø doesn’t drink coffee.
```

```
↑
```

*fully interpreted here*

In contrast, the double-negative reading is derived from the interaction of *nobody* with *not*, which cancels negation to give the interpretation in which everyone drinks coffee.

---

\(^{10}\) Given its low processing cost, one might ask why the sentence does not have a single-negative interpretation in English. The answer is that English uses an even simpler form to express that meaning: *Nobody drinks coffee.*
Nobody doesn’t drink coffee.

As illustrated here, derivation of the double-negative reading does not comply with Immediacy since the interpretation of the entire sentence has to be restructured after encountering the second negative. (Instead of meaning ‘Nobody drinks coffee,’ it has to be interpreted as ‘Everybody is a coffee-drinker.’)

On the view just outlined, the preference for Catalan-based transfer simply reflects the fact that the single-negative interpretation is the less costly option. The key intuition can be captured by adopting the following extension to the Transfer Calculus.

**The L3 Corollary**

In the case of competing ‘source languages,’ the best candidate for transfer is the operation that incurs the least cost in the L3, all other things being equal.

If this idea is on the right track, third-language acquisition is subject to the same processing pressures that shape transfer in the more widely studied L2 context – obviously a desirable state of affairs.

---

**8. Concluding Remarks**

Transfer has long occupied a prominent place in the study of second language acquisition. Indeed, it is often treated as the single most important tool for addressing the foundational puzzle of why second language learners make the particular types of errors that they do.

Crucially, though, transfer itself needs to be explained. We need to understand why it exists and why it operates in the way that it does. These priorities have been the primary concern of this chapter. The idea that I have proposed seeks to incorporate the study of transfer into a larger emergentist program devoted to processing-based explanations for the many puzzles that arise in the study of syntax and development.

If this line of inquiry is on the right track, then it makes sense to explore the possibility that transfer too is shaped by processing pressures. I have illustrated how this might work by focusing on the role of Immediacy in the operations that produce speech and derive interpretations. As I have tried to demonstrate, this approach offers an explanation for a wide range of developmental tendencies in second language acquisition.

- Why native speakers of Japanese do not prefer the Gap–Verb pattern of their first language in learning English.
• Why native English speakers reject the negated interpretation of QNPs that is preferred in their L1 when learning Korean.

• Why native Korean speakers prefer the non-negated interpretation of QNPs that is preferred in their L1 when learning English.

• Why native English speakers are reluctant to accept the distributed interpretation of indefinite NPs in Japanese.

• Why native Japanese speakers are reluctant to accept an indefinite interpretation for pro in Mandarin.

• Why bilingual speakers of Catalan and Spanish are more likely to transfer a feature of Catalan rather than Spanish to a particular pattern of negation in L3 English.

All these things happen for the same reason: as predicted by the Transfer Calculus and the two related corollaries, learners draw on the resources of a previously acquired language to reduce the cost of the operations required to learn and use a new language. The result is the complex web of effects and non-effects that we have been considering – a small fraction of the larger reality of second language acquisition.

References


