

A cross-linguistic investigation of the site of initiation in same-turn self-repair

Barbara Fox, Fay Wouk, Makoto Hayashi, Steven Fincke,
Liang Tao, Marija-Leena Sorjonen, Minna Laakso, and
Wilfrido Flores Hernandez

Introduction

Same-turn self-repair is the process by which speakers stop an utterance in progress and then abort, recast or redo that utterance. While same-turn self-repair has become a topic of great interest in the past decade, very little has been written on the question of where within a word speakers tend to initiate repair (the main exceptions being Schegloff 1979 and Jasperson 1998). And, to our knowledge, no work has been done on this question from a cross-linguistic perspective.

The goal of this paper is twofold: First, we explore existing proposals regarding where in words speakers initiate repair (what we will call the “site” of initiation) using data from our seven languages; and, second, we present and explain site of initiation data from those seven languages. Our findings suggest that there is a great deal of cross-linguistic variation with respect to favored sites of initiation but that most of the variation can be accounted for by a few simple interactional factors. This paper is the first study we are aware of which considers word length in explanations of self-repair data.

The current study is part of a larger project that has as its goal an understanding of the universal principles of self-repair and their language-specific manifestations. Prior studies have shown that the linguistic resources available to speakers of different languages shape the specific methods by which repair is accomplished (Fox *et al.* 1996; Wouk 2005; Fincke 1999; Egbert 2002; Sidnell 2007c; Karkainen *et al.*, 2007). We hope to learn more about the details of recycling and replacement repairs and how such

linguistic factors as major constituent order, frequency of ellipsis, use of prepositions or postpositions, and phonological structure of the word impact the organization of repair.

The seven languages examined in this study are: Bikol (spoken in the Philippines), Sochiapam Chinantec (spoken in Mexico), English, Finnish, Indonesian, Japanese, and Mandarin. The language sample is one of convenience, but it represents languages from six language families, three continents and three archipelagos. The languages differ in major typological features, such as basic word order, phonotactic structure, and morphological structure. In other words, the sample is genetically, typologically, and areally diverse. Below we give very brief descriptions of each language:

Bikol, a language of the Philippine branch of the Austronesian family, is verb-initial and prepositional. As in other Philippine languages, content words are multisyllabic and multimorphemic. Much of the bound morphology is prefixal, although there are infixes, suffixes, and circumfixes as well.

Sochiapam Chinantec, of the Chinantecan branch of the Oromanguan language family, is spoken in Oaxaca, Mexico. It is a verb-initial, prepositional language. Words tend to be monosyllabic and monomorphemic, but there is a small amount of derivational and inflectional morphology. Sochiapam Chinantec is a tonal language.

Finnish, in the Finno-Ugric language family, is verb-medial, with both prepositions and postpositions. It also has an extremely rich system of inflectional and derivational morphology, much of which is suffixing.

Indonesian, an Austronesian language spoken in Indonesia, is verb-medial and prepositional. Words tend to be bisyllabic, although there is a considerable amount of derivational morphology which produces longer words.

The genetic classification of Japanese remains in dispute, some treating it as essentially a language isolate, while others believe it belongs to the Altaic language family. Japanese is verb-final and postpositional. Japanese words are not divided into syllables, but rather into units called *mora*, which determine syllable weight (which, in its turn, determines stress and timing). Words tend

to consist of a large number of *mora*, and thus to be quite long in terms of their phonological structure, even though they may look no longer than a two syllable word in English. For example, “*Nippon*,” which looks like it is only two syllables, is actually four *mora*, with the double consonant and the final *n* each constituting a separate *mora*. Japanese has a rich system of bound morphology, all of which is suffixing.

Mandarin, in the Sino-Tibetan language family, is verb-medial and prepositional. It is thought of as a monosyllabic language, but modern Mandarin has such a large number of homophones that the practice of compounding has become quite prevalent, and many content words are compounds. Mandarin is also a tonal language.

Data collection

As part of a larger project on the syntax of self-repair (cf. Fox and Wouk 2003), we collected 250 instances of same-turn self-repair for each language (except English, for which we have 500 instances). The instances were taken from recordings of naturally occurring speech. The recordings were not made specifically for this project; rather, they had been collected by the individual researchers at different times and in different places, for a variety of purposes. Each instance of repair was coded for a variety of features. The coding scheme was developed for the larger project, and it included information relating to type of repair and site of initiation.

Self-repair can involve a variety of different operations, but the larger study was restricted to quite specific types. Only instances in which there was a clear syntactic relationship between the trouble source (the original utterance) and the repair were included in the collection. Thus, instances in which the speaker started an utterance, cut it off, and started a new utterance were excluded from the larger collection. In fact, the larger study is limited to cases which involve either replacement of one or more elements of the trouble source or repetition of the trouble source (or some part thereof), although either replacement or repetition (which we term “recycling”) may be accompanied by other repair operations.

Initiation in same-turn self-repair

In the present study, we further limit our focus to simple repairs, both simple recycling and simple replacement. By “simple” we mean that no other repair operation is involved in the repair. Thus, simple recycling repairs are just repetitions of words, with no additions, deletions or replacements; simple replacement repairs are instances in which a morpheme, word or phrase is replaced, without recycling; addition or deletion.¹ Since the overall study includes both simple and non-simple repetition and replacement repairs, the total number of simple repairs is less than 250 per language (or 500 for English). In this paper we examine

- 146 Indonesian repairs (117 recycling and twenty-nine replacement);
- 201 Sochiapam Chinantec repairs (185 recycling and sixteen replacement);
- 339 English repairs (285 recycling and fifty-four replacement);
- 200 Japanese repairs (147 recycling and fifty-three replacement);
- 150 Mandarin repairs (115 recycling and thirty-five replacement);
- 185 Bikol repairs (162 recycling and twenty-three replacement);
- 162 Finnish repairs (116 recycling and forty-six replacement).

As can be seen by these figures, languages differ in the proportion of simple to non-simple repairs and also in the proportion of replacement to recycling repairs, but overall simple recycling is far more frequent than simple replacement.

Methodology

There is a large body of work on self-repair in the conversation analysis (CA) framework (e.g., Jefferson 1974; Schegloff *et al.* 1977; Goodwin 1979, 1981; Schegloff 1979, 1987c, 2000c, 2002b). This work is based on qualitative analysis of individual cases in their interactional environment, or of collections of such cases. The studies are grounded in the understanding that orderliness is achieved first in single cases and that a perceived orderliness in the aggregate arises from orderliness on a single case basis. “Quantitative analysis is, in this sense, not an *alternative* to single case analysis, but

rather is built on its back" (Schegloff 1993: 102). In other words, we see similar patterns in multiple instances because speakers manage interactional contingencies with recurrent practices. It is entirely possible that orderliness at the single case basis will not lead to orderliness at the aggregate level.

In fact, Schegloff (1993) expresses serious concerns about undertaking quantitative analysis of conversational phenomena. In that paper, he raises questions about what counts as numerator and what counts as denominator. As Schegloff (1993) points out, these are not trivial questions, and in many cases the phenomenon may not be well enough understood to identify either all of the environments of relevant possible occurrence (i.e. the denominator) or all of the possible items that could count as instances of the phenomenon under investigation (i.e. the numerator). As Schegloff has suggested "We may need to defer full dress quantitative analysis in many areas until its appropriate constraints can be met" (1993: 119). For these and other reasons, CA researchers have typically avoided quantitative analysis, preferring instead to use "informal quantification." As Schegloff says:

Terminology such as occasionally or massively reports an experience or grasp of frequency, not a count; an account of an investigator's sense of frequency over the range of research experience, not in a specifically bounded body of data; a characterization of distribution fully though tacitly informed by the analytic import of what is being characterized. (Schegloff 1993: 118–119)

In this paper, however, we take findings from relevant CA work and explore their cross-linguistic implications utilizing quantitative methods. Although we are cognizant of the concerns mentioned above, we have chosen to use a quantitative approach because we are working with phenomena which are difficult for an analyst to notice at a conscious level, and to have a general grasp of, such as how often repair is initiated early in the word in a noun rather than in an article. We thus make use of quantification as a tool to help us arrive at a sense of frequency. Although this paper is full of numbers (and even statistics), we do not believe that the fact that there were, for example, X instances of repair initiated early in a noun and Y instances of repair initiated early in an article in data set P of Language Q is significant in and of itself. It only serves to

give us a sense of whether both of these syntactic types can be considered frequent sites of early repair initiation in that language.

The questions we have asked here are not just linguists' questions; they can also be participants' questions in that they involve the resources for interactional tasks such as delaying next item due or replacing a current item. For example, languages with a large number of verbal prefixes provide speakers with an easy way of delaying the verb root by recycling the prefix, while in languages which lack verbal prefixes speakers will need to find other means to delay production of the verb.

We begin our study with an exploration of the first proposal in the literature regarding site of initiation.

Site of initiation

The first proposal we will explore concerning site of initiation appears in Schegloff (1979). There, Schegloff writes, "Just post-initiation and just pre-completion of various unit types seem to be specially common loci of repair initiation. Thus, just after the start of a turn-constructional unit (e.g., a sentence) or just before its completion, *after the first sound of a word or just before its last sound*" (1979: 275).

This is just one point in a long paper covering many topics, and Schegloff does not elaborate on it nor give any examples. Nor is it clear from the paper exactly what is meant by "after the first sound of a word or just before its last sound"; after the first sound could mean after the first sound is recognizable or after the first sound is complete; just before its last sound could likewise mean before the last sound is articulated or before the last sound is complete. However, Schegloff (personal communication) has suggested that the relevant domain for just post-beginning starts after the first sound is recognizable, thus has begun to be articulated, and continues until the first sound is complete, while the relevant domain for just pre-completion begins just before the final sound is articulated, in the penultimate sound, and continues until just before the final sound is complete. He has further suggested that site of initiation of repair may depend on the type of repair involved.

The first goal of our study is to explore Schegloff's characterization of English site of initiation in the simple repairs found in the

seven languages of our sample. In order to avoid confusion of terminology, we will use the term "post-beginning" for just post-initiation of the word and "initiation" only to refer to repair initiation.

Post-beginning and pre-completion in English

We began our exploration with English. In our collection of 339 simple repairs in English, we found twenty-four simple recycling repairs (8 percent of all simple recycling repairs) and seventeen simple replacement repairs (31 percent of all simple replacement repairs) in which repair was initiated during or right after the first sound (post-beginning). Extract 1 below illustrates a simple recycling repair initiated post-beginning, in this instance in the initial vowel:

- (1) Chinese dinner
Ann: the biggest debate i- in our department. in:, at Trenton was that when we had these faculty meetings.

Extract 2 illustrates a simple replacement repair initiated post-beginning, in this instance just after the bilabial nasal closure (which could be heard to begin the word "my," replaced by *d*):

- (2) Sports
Allan: I wish I'd had m- a camera earlier today because

We further found 128 simple recycling repairs (45 percent) and twenty-two simple replacement repairs (41 percent) in which repair was initiated just before the last sound (pre-completion).

Extract 3 illustrates initiation in the penultimate sound of the demonstrative determiner "that":

- (3) SBC1 [715]

Lynne: .hN Not this time of the year, because, you know they've had a dry year, bu-
.hh usually they're too we:t, and tha- (0.7) that shoe can just pull right off;
because it's just, .hhh (.) the hoof wall is so soft.

Extract 4 shows initiation in the final sound of the pronoun *she*; the pronoun is replaced with a full noun phrase (*this girl*):

Initiation in same-turn self-repair

- (4) HG
Hyla: and she- this girl's fixed up on a da- a blind date.

It is worth noting that the vast majority of these pre-completion initiations occur in the final sound of the word, rather than prior to it.

Combining post-beginning and pre-completion frequencies for each type, we find that 152 (53 percent) of simple recycling repairs and thirty-nine (72 percent) of simple replacement repairs are in accord with Schegloff's proposal. It thus seems that for simple replacement repairs in English, Schegloff's proposed description of the common sites of repair initiation provides a rough characterization of our data. However, while post-beginning and pre-completion are both common sites of initiation of simple recycling repairs, they comprise only approximately half of these repairs. We thus wondered if other explanations might be available that would account for a greater proportion of the data. We explore other possible explanations for our findings beginning in the section 'Recognizable Completion' below. Among simple replacement repairs, on the other hand, the positions mentioned by Schegloff comprise 72 percent of all repair initiations, so it appears that Schegloff's generalization holds quite strongly for this category of repair. As Schegloff (personal communication) noted, it is to be expected that different repair types behave differently.²

Cross-linguistic comparison

Table 3.1 provides the figures for post-beginning and pre-completion sites of repair initiation for both types of repairs in all of the languages in our study. For each language, we give the percentage of all recycling repairs and replacement repairs that are initiated post-beginning and pre-completion. We also combine repairs initiated post-beginning and pre-completion, thus giving the percentage of all repairs initiated in the positions suggested by Schegloff (1979). In the discussion below, we explore the pattern of distribution of post-beginning and pre-completion repair initiation for each language, starting with Indonesian.

As can be seen from Table 3.1, in Indonesian, post-beginning is not a common site of repair initiation. Out of 146 repairs, only

Table 3.1 *Post-beginning and pre-completion repair initiation*

		Post-beginning	Pre-completion	Combined	Total repairs
English	Recycling	24 (8%)	128 (45%)	152 (53%)	285
	Replacement	17 (31%)	22 (41%)	39 (72%)	54
Indonesian	Recycling	6 (5%)	51 (35%)	57 (39%)	117
	Replacement	0 (0%)	8 (28%)	8 (28%)	29
Sochiapam Chinantec	Recycling	2 (1%)	22 (12%)	24 (13%)	185
Mandarin	Replacement	1 (6%)	7 (44%)	8 (50%)	16
	Recycling	5 (4%)	33 (29%)	38 (33%)	115
Bikol	Replacement	2 (6%)	6 (17%)	8 (23%)	35
	Recycling	8 (5%)	37 (23%)	45 (28%)	162
Finnish	Replacement	1 (4%)	11 (48%)	12 (52%)	23
	Recycling	21 (18%)	19 (16%)	40 (34%)	116
Japanese	Replacement	13 (28%)	7 (15%)	20 (43%)	46
	Recycling	30 (20%)	39 (27%)	69 (47%)	147
	Replacement	2 (4%)	17 (32%)	19 (36%)	53

six simple recycling repairs (5 percent) and no simple replacement repairs were initiated immediately post-beginning. Pre-completion, on the other hand, does seem to be a moderately frequent site, as fifty-one simple recycling repairs (35 percent) and eight simple replacement repairs (28 percent) are initiated pre-completion. Combining post-beginning and pre-completion frequencies for each type, we find that fifty-seven (39 percent) simple recycling repairs and eight (28 percent) simple replacement repairs are in accord with Schegloff's proposal.

In Sochiapam Chinantec and Mandarin, as in Indonesian, post-beginning is not a common site of repair initiation; in Sochiapam Chinantec only two simple recycling repairs and one simple replacement repair are initiated post-beginning, while in Mandarin two (6 percent) simple replacement repairs and five (4 percent) simple recycling repairs were initiated post-beginning. The two languages differ from each other in the degree to which speakers seem to orient to pre-completion. In Sochiapam Chinantec, pre-completion is a proportionally more frequent site for replacement repair initiation – being found in seven cases (44 percent) – than for recycling repair initiation, where it is found in twenty-two cases (12 percent). In Mandarin, it is a more frequent site for recycling repair initiation – being found in thirty-three cases (29 percent) – than for replacement repairs, where it is found in six cases (17 percent). Combining post-beginning and pre-completion frequencies for each type, we find that in Sochiapam Chinantec twenty-four (13 percent) simple recycling repairs and eight (50 percent) simple replacement repairs are in accord with Schegloff's proposal, while in Mandarin thirty-eight (33 percent) simple recycling repairs and eight (23 percent) simple replacement repairs are.

In Bikol as well, post-beginning is not a frequent site of repair initiation; only nine (5 percent) simple repairs were initiated immediately post-beginning, eight simple recycling repairs and one simple replacement repair. Among simple recycling repairs, thirty-seven (23 percent) were initiated just pre-completion, while among simple replacement repairs, eleven (48 percent) were initiated pre-completion. Combining post-beginning and pre-completion frequencies for each type, we find that forty-five (28 percent)

simple recycling repairs and twelve (52 percent) simple replacement repairs are in accord with Schegloff's proposal.

In Finnish, as in English, post-beginning is a moderately frequent site of repair initiation, particularly for replacement repairs. We found that thirty-four (21 percent) Finnish simple repairs, twenty-one (18 percent) recycling repairs and thirteen (28 percent) replacement repairs, were initiated immediately post-beginning. Pre-completion is also used with some frequency, although less frequently than post-beginning, as a site of repair initiation in Finnish. Among simple recycling repairs, nineteen (16 percent) repairs were initiated just pre-completion. Among simple replacement repairs, seven (15 percent) were initiated just pre-completion. Combining post-beginning and pre-completion frequencies for each type, we find that forty (34 percent) simple recycling repairs and twenty-six (43 percent) simple replacement repairs are in accord with Schegloff's proposal.

In Japanese, repair is initiated in post-beginning position with moderate frequency for recycling, but not for replacement, as thirty-two (16 percent) repairs – thirty (20 percent) simple recycling repairs and two (4 percent) simple replacement repairs – were initiated immediately post-beginning. Pre-completion initiation is moderately frequent for both repair types. Among simple recycling repairs, thirty-nine (27 percent) were initiated just pre-completion. Among simple replacement repairs, seventeen (32 percent) were initiated just pre-completion. Combining post-beginning and pre-completion frequencies for each type, we find that sixty-nine (47 percent) simple recycling repairs and nineteen (36 percent) simple replacement repairs are in accord with Schegloff's proposal.

Summary of cross-linguistic comparison

What we have found for simple repairs in our seven languages, then, is that Schegloff's proposed description matches English the best and is a particularly useful generalization about site of initiation in simple replacement repairs. However, the pattern is clearly not universal, and most of the matching in our other languages is accounted for by pre-completion, not by post-beginning. In fact, in several languages (Indonesian, Bikol, Sochiapam Chinantec, and

Mandarin) speakers do not orient toward post-beginning at all, while in others speakers show an orientation toward it only for one type of repair (simple replacement repairs in English and simple recycling repairs in Japanese), and only in Finnish do speakers seem to orient toward it for both types. In all languages, speakers showed some orientation toward pre-completion, but varied greatly in how much. Furthermore, while speakers of all of the languages in our study show an orientation toward at least one of these two sites of repair initiation, only in English are these sites particularly common.

Recognizable completion

We have seen that post-beginning is far less relevant cross-linguistically than pre-completion, as illustrated in Table 3.1. Additionally, the majority of the pre-completion repair initiations in all the languages of our study occurred in the final segment, rather than just before it. Furthermore, for several of the languages in our study (Bikol, Finnish, Indonesian, and English) between one quarter and one third of recycling repairs were initiated after the word had been completely articulated, and for Mandarin and Sochiapam Chinantec the majority of recycling repairs were initiated in this location. Only in Japanese was this site rarely used. It would seem, then, that for most languages recycling repairs tend to be initiated either in the final sound of the word or after it. It might be useful to consider why this is the case.

We will refer to initiation in or after the last sound of the word as initiation at recognizable completion. In four of the languages in our study (English, Mandarin, Sochiapam Chinantec, and Indonesian), speakers favor initiating repair at recognizable completion in recycling repairs. However, not all languages in the study showed this orientation; we demonstrate in 'Cross-Linguistic Comparison' below (as illustrated in Tables 3.4, 3.5, and 3.6), that speakers of Finnish, Bikol, and Japanese seem to have no special preference for sites of initiation after the word is recognizably complete. We will consider a number of factors below in explaining the difference between these two groups of languages.

Given the work that the concept "recognizable completion" seems to do for speakers of at least four of our languages, it seems promising to broaden Schegloff's categories of post-beginning and pre-completion to "before recognizable completion" and "recognizable completion" as possible sites of repair initiation oriented to by speakers. Recognizable completion includes most cases where repair is initiated pre-completion and also those cases where repair is initiated after the word is complete.

Initiations in the penultimate sound (part of Schegloff's original characterization of pre-completion) would be excluded, but there turned out to be very few of them in our data, in any language. All other sites of repair initiation would be considered a single category, "before recognizable completion," which would contrast with "recognizable completion."³

It has long been known (Sacks *et al.* 1974) that participants orient to recognizable completion in their monitoring of the turn-so-far for transition relevance; however, it could be argued that speakers do not orient to the recognizable completion of a word but rather only to the recognizable completion of a turn constructional unit (TCU), and therefore that the relevant characterization of within-word repair sites would be something else. Another possible characterization, for example, could be that speakers tend to initiate repair late in a word, without regard to projection of completion. Although little research has been done on speakers' and recipients' monitoring of word production (but see Jasperson 1998; Jefferson 1974), our findings suggest that "recognizable completion" better captures the phenomenon. If lateness were the appropriate characterization, we would expect to find speakers frequently initiating repair in the position just prior to the final sound, as well as in that sound and after it, since just before the final sound is also late. However, we do not find this. What we find is that speakers overwhelmingly orient to recognizable completion in initiating repair. This is a fact that needs to be captured.

So, how is it that speakers and recipients orient to the completion-relevant trajectory of words, in addition to TCUs? We suggest here that while recipients monitor the course of an utterance for where it might come to possible completion, they are also monitoring the unfolding of an utterance for its relevance to other forms of co-participation (e.g., head nod, gaze organization,

assessment-relevant activity) (cf. Ford *et al.* 1996), and for the moment-by-moment, word-by-word, shifting course of the construction of the TCU (cf. Ford *et al.* 1996). Each word shapes the trajectory of the TCU, either renewing the current trajectory or changing it. It thus makes sense that speakers and recipients should orient to the word as a unit of production, and to its completion.

Furthermore, although it may not be the case for English, in many languages a word's identity is not certain until it is complete. For example, in Japanese and Finnish, which make use of a large number of inflectional and derivational suffixes, a word's syntactic class, and therefore the trajectory of the turn so far, can remain indeterminate until the end.

In the remainder of the study, then, we will make use of the concepts "recognizable completion" and "before recognizable completion" as within-word sites of repair initiation.

And how do replacement repairs pattern with regard to recognizable completion? It appears that they are treated quite differently from recycling repairs, at least in some languages in our survey. For example, recognizably complete initiations comprise only 35 percent of all simple replacement repairs in English. This interesting difference between recycling and replacement in the use of the category recognizably complete will be discussed further below. We got our inspiration about how to think about this difference between the two repair types from the only major work on site of repair initiation, Jasperson (1998). We turn now to a summary of that work.

Site of initiation and repair type

Jasperson first introduces a distinction between two different types of cut-offs, word-disruptive cut-off and word-preserving cut-off, a distinction he clarifies as follows: "Word-preserving cut-off is cut-off which occurs upon recognizable completion of the word, so that by the time cut-off occurs, a word has been produced intact. In contrast, word-disruptive cut-off is cut-off which occurs at any point before recognizable completion of the word-in-progress, resulting in a 'word fragment'" (1998: 140).

Repair that is initiated before a word is recognizably complete is thus word-disruptive. During the final sound of a word, speakers

can produce either word-disruptive or word-preserving cut-off: In word-preserving cut-off, the cut-off is done softly, and the final segment is not shortened or perturbed in any way, while in word-disruptive cut-off the cut-off is louder, the final segment is typically shortened, and glotalization may begin during a vowel preceding the final segment in a consonant-final word.

Jasperson then goes on to suggest a correlation between these cut-off types and the type of repair they construct: "The orientation of word-disruptive cut-off is typically retrospective, insofar as it routinely initiates repair which operates to change some part of the turn in progress (cf. Schegloff 1979). [...] Word-preserving cut-off on the other hand does not seem to have this retrospective orientation" (1998: 77).

Jasperson's findings thus suggest that repair initiated before a word is recognizably complete tends to be engaged with retrospective work, while repair at recognizable completion is not engaged with retrospective repair.

In our study we included a wide variety of initiation types, not just cut-off, and did not make use of phonetic distinctions between different types of cut-offs. Rather, we focused on where in the word initiation took place. This decision was largely practical; many of the data sets being used were not of sufficiently high sound quality to make detailed phonetic distinctions, nor did the researchers all possess the auditory acuity for accurate coding. Therefore, we could not test Jasperson's proposal against other languages. However, Jasperson's findings encouraged us to look for other correlations between repair type and repair initiation. In particular, we looked at whether retrospective vs. non-retrospective repair might correlate with initiation when the word is recognizably complete, or earlier in the word, before it is recognizably complete. We therefore looked to see whether the difference between recycling and replacement noted above could be explained in terms of sensitivity to recognizable completion. Based on Jasperson's study, we hypothesized that initiation before the word is recognizably complete is associated with repairs that change an element in the preceding talk (for example by replacing it), while initiation when the word is recognizably complete is associated with repairs that operate on upcoming talk, for example by delaying next item due (by recycling current word). The following section tests this hypothesis for English.

Table 3.2 *Recognizable completion and repair initiation in English*

	Not recognizably complete	Recognizably complete	Total
Recycle	58 (20%)	227 (80%)	285
Replace	35 (65%)	19 (35%)	54
Total	93 (27%)	246 (73%)	339

Chi-square = 45.09, d.f. = 1, $p = .0000$

Repair type and site of initiation in English

Our English data support this claim strongly: 65 percent of all simple replacement repairs are initiated before the word is recognizably complete, while only 20 percent of simple recycling repairs are (see Table 3.2 below). This difference is significant ($p = .0000$), and both repair types contribute to the significance. It thus appears that in English there is a strong association between site of initiation and repair type, even when a wide range of types of repair initiation are included.

Before examining the patterns in our other languages, it is worth exploring possible explanations for this association in English. One obvious explanation is the function of each repair type. Recycling can be used to elicit gaze from recipients or to treat the start of an utterance not accompanied by recipient gaze as defective (Goodwin 1981); it can also be used to delay the next item due, either in a word search or as an indication that the speaker is choosing between alternatives (Jefferson 1974); further, recycling is found in overlaps, as a device to produce talk in the clear (Schegloff 1987c). While all of these functions are found in our English data, we have found that recycling in English tends to be used to repeat function words, which we hypothesize is most often done in order to delay the production of the next content word due (80 percent of all simple recycling in our English data repeat function words). Because their purpose is exactly to delay production of next item due, there is no hurry in initiating the repair. Replacement, on the other hand, tends to be used to replace content words (61 percent of simple replacement repairs in

English replace content words) and may occur in cases where an inappropriate word or pronunciation has been produced. Jefferson (1974) provides some insight into this phenomenon; she suggests that speakers make use of the format word₁ + hesitation + word₂ to replace one word with another, and that speakers can employ this device to produce just enough of a word for the word to be recognizable, at least retrospectively, thereby allowing the speaker to produce an inappropriate word without being interactionally accountable for it.

Or, for example, one can propose 'I am not like this but am talking by reference to the fact that *you* are' by finding ways to show that the terms one produces are not the terms which first come to mind. (Jefferson 1974: 193)

One means of achieving that display may be the production of just enough error to convey one's habitual terminology without inheriting complaints from its recipient (i.e. not having "officially" produced the word in question) and then correcting it with a term which can be seen as selected by reference to one's situation or recipient. (Jefferson 1974: 193)

As Jefferson indicates, in such cases there may be pressure to initiate repair as early as possible, to produce a word without being accountable for its production. Of course, not all replacements are so interactionally significant, but speakers may nevertheless prefer to minimize the production of an errorful or inappropriate word (cf. Jasperson 1998).⁴ However, the pressure may be less in these cases than in the type of instances that Jefferson examines, thus allowing for initiation after recognizable completion.⁵

Our explanation for differences in preferred site of repair initiation for recycling and replacement repairs thus appears to rest on what might be universal motivations. But do other languages actually manifest the same patterns as English? The following section examines how the other languages pattern.

Cross-linguistic comparison

If we look at Indonesian, the results appear promising. Indonesian patterns much like English, as shown in Table 3.3 below, with the majority (74 percent) of recycling repairs being initiated when the word is recognizably complete, while only a minority (28 percent) of the replacements are initiated in that site. As in English, the

Table 3.3 *Recognizable completion and repair initiation in Indonesian*

	Not recognizably complete	Recognizably complete	Total
Recycle	30 (26%)	87 (74%)	117
Replace	21 (72%)	8 (28%)	29
Total	51 (35%)	95 (65%)	146

Chi-square = 22.37, d.f. = 1, $p = .0000$

Table 3.4 *Recognizable completion and repair initiation in Finnish*

	Not recognizably complete	Recognizably complete	Total
Recycle	52 (45%)	64 (55%)	116
Replace	35 (76%)	11 (24%)	46
Total	87 (54%)	75 (46%)	162

Chi-square = 12.95, d.f. = 1, $p = .0003$

difference is highly significant ($p = .0000$), and both repair types contribute to the significance.

Finnish is somewhat similar, with speakers strongly favoring initiations before the word is recognizably complete for replacement repairs; recycling repairs, however, are more evenly distributed between initiations before the word is recognizably complete and when the word is recognizably complete, as shown in Table 3.4. Note that while the overall distribution in Table 3.4 is significant ($p = .0003$), the significance comes entirely from the Replace row. Recycling repairs are actually distributed evenly with respect to repair initiation type (chi-square = 1.24; $p = 0.27$).

However, when we examine Bikol (Table 3.5) and Japanese (Table 3.6), we can see immediately that the pattern as manifested by English and Indonesian is not universal. For Japanese ($p = .5649$) and Bikol ($p = .8428$) initiation type does not pattern significantly according to repair type.

From Table 3.5 we can see that in Bikol speakers have no preference for either site of initiation, with both recycling and

Table 3.5 *Recognizable completion and repair initiation in Bikol*

	Not recognizably complete	Recognizably complete	Total
Recycle	88 (54%)	74 (48%)	162
Replace	13 (56%)	10 (44%)	23
Total	101 (55%)	84 (45%)	185

Chi-square = 0.04, d.f. = 1, p = .8428

Table 3.6 *Recognizable completion and repair initiation in Japanese*

	Not recognizably complete	Recognizably complete	Total
Recycle	106 (72%)	41 (28%)	147
Replace	36 (68%)	17 (32%)	53
Total	142 (71%)	58 (29%)	200

Chi-square = 0.33, d.f. = 1, p = .5649

replacement repairs being initiated before recognizable completion roughly 55 percent of the time; Table 3.6 suggests that speakers of Japanese favor repair before recognizable completion, with replacement repairs initiated before recognizable completion 68 percent of the time and recycling repairs initiated before recognizable completion 72 percent of the time. The difference between the two languages is highly significant (chi-square = 11.11, d.f. = 1, p = .0009, chi-square calculated on the totals for not recognizably complete and recognizably complete for each language). In fact, Japanese proved to be the only language in our sample where speakers actually favor initiation before recognizable completion for recycling repairs.

Turning now to Sochiapam Chinantec (Table 3.7) and Mandarin (Table 3.8), we see yet another pattern, namely that speakers strongly favor initiation when the word is recognizably complete for both recycling and replacement repairs, with no significant difference between the two types of repair (Sochiapam Chinantec p = .2385, Mandarin p = .4006).

Table 3.7 *Recognizable completion and repair initiation in Sochiapam Chinantec*

	Not recognizably complete	Recognizably complete	Total
Recycle	26 (14%)	159 (86%)	185
Replace	4 (25%)	12 (75%)	16
Total	30 (15%)	171 (85%)	201

Chi-square = 1.39, d.f. = 1, p = .2385

Table 3.8 *Recognizable completion and repair initiation in Mandarin*

	Not recognizably complete	Recognizably complete	Total
Recycle	12 (10%)	103 (90%)	115
Replace	2 (6%)	33 (94%)	35
Total	14 (9%)	136 (91%)	150

Chi-square = 0.71, d.f. = 1, p = .4006

There is no significant difference between Mandarin and Sochiapam Chinantec with respect to initiation type of simple recycling repairs (chi-square = .840, d.f. = 1, p = 3595, chi-square calculated on the values for recognizably complete and not recognizably complete for recycling repairs in the two languages). There are too many cells with values below 5 to attempt a statistical comparison of the two languages with respect to simple replacement repairs, but the pattern is clear: In both languages speakers strongly favor initiation after recognizable completion for replacement repairs.

Summary

With a sample of seven languages we have seen five of the logically possible patterns for site of initiation and repair type:

1. no difference, with speakers favoring initiation before the word is recognizably complete for both types of repairs (Japanese);

2. no difference, with speakers favoring initiation when the word is recognizably complete for both types of repairs (Mandarin, Sochiapam Chinantec);
3. no difference, with speakers favoring neither type of initiation (Bikol);
4. speakers favoring initiation when the word is recognizably complete for recycling repairs and initiation before the word is recognizably complete for replacement repairs (English, Indonesian);
5. speakers favoring initiation before the word is recognizably complete for replacement repairs, and showing no preference for recycling repairs (Finnish).

It is difficult to argue for a universal pattern when so many of the logical possibilities are found in such a small sample of languages.

Given our discussion of the interactional functions of recycling and replacement repairs in "Monosyllabic Words and Recognizable Completion," this is a surprising finding. We would have expected that data from other languages would give evidence of speakers orienting to the interactional pressures we discussed, and at least on the surface they appear not to.

Nonetheless, we believe that there are underlying universal patterns at work here. Recycling tends to be initiated after recognizable completion, as it is frequently employed to delay the next content word due. The only exception to this is Japanese, where recycling repairs are generally initiated prior to recognizable completion and where function words generally follow content words. Replacement tends to be initiated prior to recognizable completion to reduce accountability for inappropriate words. However, word length, that is the number of syllables in a word,⁶ and syntactic class of the word in which repair is initiated, are also important factors, and their influence can at least partially explain our surprising finding. In the next section, we explore in detail the underlying universal patterns relating to site of initiation, word length and syntactic class.

Word length organization and site of initiation

The languages in our sample vary in their morphological structure and especially in the length of words involved in repair: Mandarin,

Table 3.9 Site of initiation of simple recycling repair initiated in monosyllabic words

	Not recognizably complete	Recognizably complete	Total
Sochiapam	4 (3%)	137 (97%)	141
Chinantec	1 (4%)	24 (96%)	25
Indonesian	5 (6%)	86 (94%)	91
Mandarin	7 (12%)	43 (88%)	50
Bikol	39 (16%)	209 (84%)	248
English	8 (21%)	30 (79%)	38

Chinantec, and English speakers tend to initiate repair in monosyllabic words;⁷ Indonesian speakers most often initiate repair in bisyllabic words; and Bikol, Japanese, and Finnish speakers tend to initiate repair in multisyllabic words of three or more syllables.⁸ So, one obvious question to ask of the data is whether word length might play a part in site of initiation for the two types of repair. We explored this question by looking at site of initiation for words of varying length in each language. Our findings suggest that word length is, indeed, a factor in site of initiation and that cross-linguistic patterns underlie that outcome.

Monosyllabic words and recognizable completion

The first positive result that came from our study was that in all of the languages speakers favor "recognizably complete" initiation for recycling repair initiated in monosyllabic words. Tables 3.9 and 3.10 below provide the data for all seven languages.⁹

From Table 3.9 it is clear that in all the languages in the sample speakers favor initiation when the word is recognizably complete for recycling repairs in monosyllabic words, although with some variation in the strength of that preference. In four of the languages (Japanese, Sochiapam Chinantec, Indonesian, and Mandarin), speakers show a very strong preference, with well under 10 percent initiation before the word is recognizably complete. In the remaining three (Bikol, Finnish, and English) speakers show a strong preference, with percentages of initiation before the word is recognizably complete above 10 percent but below 20 percent.

Table 3.10 *Site of initiation of simple replacement repair initiated in monosyllabic words*

	Not recognizable complete	Recognizably complete	Total
Indonesian	0 (0%)	0 (0%)	0
Bikol	0 (0%)	3 (100%)	3
Finnish	0 (0%)	4 (100%)	4
Japanese	0 (0%)	5 (100%)	5
Mandarin	2 (6%)	28 (94%)	30
Sochiapam	1 (8%)	11 (92%)	12
Chinanteec English	17 (50%)	17 (50%)	34

If we consider site of initiation in replacement repairs in monosyllabic words, we find a similar pattern for all languages except English (and Indonesian, which had no instances of replacement repair initiated in monosyllabic words). Note that in some cases it is not possible to determine the number of syllables of a word that gets replaced, since its lexical identity is not always clear. This proved to be a problem particularly for the English and Finnish data sets. Therefore, in this section of the paper, the data reported for English and Finnish includes only cases where word length could be determined (forty-six cases for English and twenty-six cases for Finnish).

As shown in Table 3.10, in several of the languages in the study, those where speakers regularly use a large number of longer words in ordinary conversation, speakers rarely initiate simple replacement repairs in monosyllabic words. However, when they do, they follow a consistent pattern. With the exception of English (and Indonesian), the preference for initiation when the word is recognizable complete in monosyllabic word replacement ranges between 90 percent and 100 percent. Frequencies are too low for reliable statistical analysis, but the difference between English and all the other languages in the study is noticeable, and we expect that larger data sets would bear it out. If we sum the results from the languages besides English and Indonesian, there are three instances of initiation prior to recognizable completion and fifty instances after recognizable completion. It is possible to compare these totals with English, and the result is highly significant ($\chi^2 = 23.47$,

d.f. = 1, $p = .0000$). The difference between English and all the other languages in the study in this regard will be taken up in "Explaining English" below.

We have thus uncovered our first universal pattern: Recycling repair is initiated when the word is recognizable complete in monosyllabic words. Replacement repairs also tend to be initiated when the word is recognizable complete in monosyllabic words, except in English. This is surprising, as our understanding of the interactional functions of replacements would lead us to expect that other languages would be more like English in initiating repair more often before recognizable completion. However, we will argue later ("Explaining English") that there are reasons why we see this difference, reasons that are compatible with our understanding of interactional functions of repair. We will suggest that monosyllabic words are typically function words in many languages, and therefore pattern differently than one might expect (for content words), as they are less likely to be interactionally delicate.

Multisyllabic words and recognizable completion

If we compare the pattern of repair initiation in monosyllabic words to the patterns of repair initiation in multisyllabic words, we come to another important finding. Consider Table 3.11, which provides the frequencies for site of initiation in recycling repair in words of three syllables or more:

We can notice first that the numbers for Sochiapam Chinanteec, English, and Mandarin are very small, because in those languages speakers tend to use mostly words of one and two syllables in ordinary conversation. In fact, the totals for Mandarin and Sochiapam Chinanteec are so low that one really cannot talk about tendencies in any meaningful way. Therefore, the discussion here will be restricted to the five languages that have ten or more multisyllabic words in which repair is initiated. In all languages except Indonesian (46 percent), speakers show a preference for initiation before the word is recognizable complete in these longer words, with Bikol, Finnish, and Japanese speakers showing the greatest preference for such initiation (between 80 percent and 90 percent). If we compare Table 3.11 to Table 3.10, we see that in each of the languages for which we have sufficient numbers to make a

comparison, speakers show a dramatic difference in preferred site of repair initiation between multisyllabic words and monosyllabic words, with a higher frequency of initiations before the word is recognizably complete in multisyllabic words.

Table 3.12 provides a similar set of figures for replacement repairs in multisyllabic words. For each language, we give the numbers for site of initiation in replacement repair in words of three syllables or more.

Table 3.11 *Site of initiation in recycling repairs in words of three or more syllables*

	Not recognizably complete	Recognizably complete	Total
Mandarin	1 (33%)	2 (67%)	3
Indonesian	12 (46%)	14 (54%)	26
English	7 (70%)	3 (30%)	10
Finnish	24 (77%)	7 (23%)	31
Japanese	80 (87%)	12 (13%)	92
Bikol	61 (91%)	6 (9%)	67
Sochiapam	1 (100%)	0 (0%)	1
Chinantec			

Japanese/Finnish: chi-square = 1.60, d.f. = 1, $p = 0.2$
 Finnish/English: chi-square = 0.23, d.f. = 1, $p = .63$
 English/Japanese: chi-square = 2.07, d.f. = 1, $p = .1505$
 Japanese/Bikol: chi-square = 0.65, d.f. = 1, $p = .4218$
 Indonesian/Bikol: chi-square = 22.36, d.f. = 1, $p = .0000$

Table 3.12 *Site of initiation in replacement repairs in words of three or more syllables*

	Not recognizably complete	Recognizably complete	Total
Bikol	8 (62%)	5 (38%)	13
Finnish	7 (64%)	4 (36%)	11
Indonesian	10 (83%)	2 (17%)	12
Finnish	27 (87%)	4 (13%)	31
Japanese	31 (89%)	4 (11%)	35
English	4 (100%)	0 (0%)	4
Sochiapam	0 (0%)	0 (0%)	0
Chinantec			
Mandarin	0 (0%)	1 (100%)	1

When we consider site of replacement repairs, we find that in most of the languages in our study speakers show a preference for initiation before the word is recognizably complete in multisyllabic words, as illustrated in Table 3.12. Sochiapam Chinantec and Mandarin lack significant numbers of multisyllabic words and are thus not relevant to this question. In all the remaining languages, speakers most frequently initiate repair in multisyllabic words prior to recognizable completion. It is likely that this finding relates to the fact that multisyllabic words are generally content words and thus more likely to be interactionally delicate or potentially inappropriate.¹⁰

We now have two important findings: repair is initiated when the word is recognizably complete in monosyllabic words and is initiated before the word is recognizably complete more frequently in multisyllabic words than in monosyllabic ones. We are thus in a position to suggest that if speakers of a language favor monosyllabic words as the site of repair initiation, they will tend to initiate repair when the word is recognizably complete. If speakers of a language favor multisyllabic words, they will tend to initiate repair before the word is recognizably complete. If we look at the languages through this perspective, we will see that the tendencies do hold up, although there are some interesting exceptions.

Word length, repair type and recognizable completion

In this section we further explore the correlation between repair type and word length in the languages of our study. Tables 3.12 and 3.13 give the percentage of monosyllabic and multisyllabic words in both repairs types for each language. Table 3.12 provides the results for monosyllabic words.

From Table 3.13, we can see that in Sochiapam Chinantec, English, and Mandarin, speakers typically initiate repair in one-syllable words, in both kinds of repair. In Bikol, Finnish, and Indonesian, about one-third to one-fifth of recycling repairs are initiated in one-syllable words, and in Bikol and Finnish replacement repairs, an even lower proportion is found (Indonesian has no monosyllabic words in which replacement repairs are initiated). Japanese is unique in having a very low rate of one-mora words in which recycling repairs are initiated; however, this may

Table 3.13 Frequency of monosyllabic words for each repair type

	Recycle		Replace	
	N	%	N	
Japanese	8	5	5	9
Indonesian	25	21	0	0
Bikol	50	30	3	13
Finnish	37	32	4	9
Sochiapam	141	76	12	75
Chinanteec				
Mandarin	91	79	30	86
English	284	87	34	74

Table 3.14 Frequency of multisyllabic words for each repair type

	Recycle		Replace	
	N	%	N	
Sochiapam	1	1	0	0
Chinanteec				
Mandarin	3	3	1	3
English	10	4	4	9
Indonesian	26	22	12	41
Finnish	31	27	31	67
Bikol	67	41	13	56
Japanese	92	63	35	66

relate to the fact that one-*mora* words are somewhat rare in Japanese. Moreover, one-*mora* words are mostly postpositions (e.g., *ni*, *ga*, *de*, *no*), and because postpositions follow rather than precede their nouns, they are generally not used to delay next content word due. They thus have a limited functionality in recycling repairs.

Turning now to multisyllabic words, we see a picture that is complementary to the one in Table 3.13. From Table 3.14 we can see that in Sochiapam, Chinanteec, Mandarin, and English speakers do not often initiate either recycling or replacement repair in words of three or more syllables. Indonesian and Finnish have a somewhat higher percentage of recycling repairs initiated in long words and a

substantially higher percentage of replacement repairs; the increase is particularly striking for Finnish, which actually has the highest percentage of multisyllabic words in which repair is initiated in the languages of our sample. Japanese and Bikol have a fairly high percentage of repair initiation in longer words for both recycle and replacement repairs.

If we take these findings back to our original question of site of initiation, it seems that we can now explain some of the basic patterns. Sochiapam, Chinanteec and Mandarin speakers initiate repair mainly in monosyllabic words, and, since there appears to be a universal tendency toward initiation when the word is recognizable complete when repair is initiated in monosyllabic words, these two languages show an overwhelming preference for initiation at recognizable completion for both kinds of repair.

Finnish and Bikol have a moderate number of monosyllabic words in which recycling repairs are initiated and a low number of monosyllabic words in which replacement repairs are initiated. Finnish has a moderate rate of multisyllabic words in which recycling repairs are initiated and a quite high rate of multisyllabic words in which replacement repairs are initiated. Bikol has a moderately high frequency of multisyllabic words in which repair is initiated for both types of repair. We would thus expect an increased frequency of initiation before the word is recognizably complete, when compared to Sochiapam, Chinanteec and Mandarin, because repairs in multisyllabic words tend to be initiated before the word is recognizably complete, and both Bikol and Finnish have a greater proportion of multisyllabic words than do Sochiapam, Chinanteec and Mandarin. Both languages match this prediction in having an increase in initiation before recognizable completion for both repair types.

Japanese is unusual in having an extremely low rate of monomoraic words as sites of initiation for both kinds of repair, and a fairly high rate of multimoraic words as sites of initiation for both kinds of repair. We would thus expect a strong preference for early initiation for both kinds of repair, since repairs in longer words tend to be initiated before recognizable completion, and this is what we find.

In Indonesian, speakers initiate recycling repairs in monosyllabic words with only moderate frequency, and no replacement repairs at all are initiated in monosyllabic words. Recycling repairs are

initiated in multisyllabic words with a moderately high frequency. Indonesian thus clearly does not pattern with English, Mandarin, and Sochiapam Chinantec in terms of the length of word in which repair is most often initiated. Nor does it pattern with Japanese, as the frequency of repair initiation in multisyllabic words is considerably lower in Indonesian than in Japanese. Indonesian is more like Finnish or Bikol, as all three languages have only moderate numbers of monosyllabic words in which repair is initiated and the frequency of multisyllabic words in which repair is initiated in Indonesian is closest to their frequency in Finnish for recycling repairs and in Bikol for replacement repairs. Indonesian might thus be expected to pattern with Finnish and Bikol and to have an increase in initiation before recognizable completion for both repair types compared with Mandarin and Sochiapam Chinantec. However, the Indonesian data does not show the same preference for site of initiation for the two repair types. Instead, Indonesian speakers prefer initiation after recognizable completion for recycling repairs and initiation prior to recognizable completion for replacement repairs.

Indonesian speakers' preference for initiation prior to recognizable completion in replacement repairs fits our prediction, since longer words show a preference for initiation before recognizable completion, but their preference for initiation after recognizable completion for recycling repairs does not (they should show a greater preference for early initiation even in recycling repairs because repairs are initiated in multisyllabic words with moderate frequency). Indonesian thus appears to be an exception to our pattern; below we suggest that there is a competing cross-linguistic factor at work in Indonesian which produces this result.

English is another exception to the findings regarding word length. A preponderance of English recycling repairs are initiated in monosyllabic words, so we would expect English speakers to favor initiation after recognizable completion for recycling repairs, which it does. However, as noted above, English speakers favor initiation prior to recognizable completion for replacement repairs, even in monosyllabic words. We have already suggested that this is a response to interactional pressures, and we will further explore the difference between English and most of the other languages in our corpus below.

Table 3.15 *Site of initiation of recycling repairs in bisyllabic words*

	Not recognizably complete	Recognizably complete	Total
Indonesian	17 (26%)	49 (74%)	66
Mandarin	6 (29%)	15 (69%)	21
Finnish	20 (43%)	27 (57%)	47
Bikol	20 (44%)	25 (56%)	45
English	12 (44%)	15 (56%)	27
Sochiapam Chinantec	21 (49%)	22 (51%)	43
Japanese	26 (55%)	21 (45%)	47

It appears that only five languages (Sochiapam Chinantec, Mandarin, Finnish, Bikol, and Japanese) fit the predictions from our general findings about word length and site of repair initiation, while two (Indonesian and English) do not. We argue below that there are principled reasons for these two exceptions.

Explaining Indonesian: Longer words and recognizable completion

We have shown that Indonesian is exceptional in that, according to the word-length criterion, speakers should show a greater preference for initiation prior to recognizable completion in recycling repairs than they do. We propose to explain this exception in terms of word length and syntactic class. The careful reader will no doubt have noticed by now that we have discussed monosyllabic words and multisyllabic words but have ignored bisyllabic ones. The time has come to rectify this matter. Table 3.15 presents the results for bisyllabic words in recycling repairs. Due to the small number of replacement repairs initiated in bisyllabic words, we will not discuss them in this paper.

We find two different patterns in our data. In Indonesian and Mandarin, speakers prefer initiation after recognizable completion in bisyllabic words. To put it another way, they seem to treat bisyllabic words much as they treat monosyllabic ones. In the remaining languages, speakers do not show a preference for site of initiation in bisyllabic words, suggesting that for speakers of those languages

bisyllabic words are treated as intermediate between monosyllables, in which repair is typically initiated after recognizable completion and multisyllabic words, in which repair is typically initiated prior to recognizable completion. The preponderance of initiations after recognizable completion in bisyllabic words in Mandarin is entirely in line with the general pattern of initiation after recognizable completion in Mandarin.¹¹

Indonesian is unusual in our sample in that repair is most often initiated in bisyllabic words, rather than monosyllabic or multisyllabic words (see Table 3.16). More than half (sixty-six or 56 percent) of simple recycling repairs in Indonesian are initiated in bisyllabic words, whereas the proportion for most of the other languages ranges from one fifth to one third, and for English is far lower (twenty-seven or 9 percent). Only Finnish approaches Indonesian in the frequency with which repair is initiated in bisyllabic words, with forty-seven (41 percent), but even in this case the difference between the two languages is statistically significant (chi-square 5.89, d.f. = 1, $p = .0152$). The fact that Indonesian does not fit our prediction, and that speakers of Indonesian should prefer repair initiation prior to recognizable completion more frequently than they do in recycling repairs appears to be traceable to the high frequency with which repair is initiated in bisyllabic words.

Why should Indonesian speakers so frequently initiate recycling repairs in bisyllabic words? We believe the key lies in the nature of function words in Indonesian: Most function words in Indonesian are bisyllabic (including pronouns, prepositions, and determiners). Indonesian is like English in that function words precede content

Table 3.16 *Percentage of recycling repairs initiated in bisyllabic words*

	Bisyllabic	All recycling
Indonesian	66 (56%)	117
Finnish	47 (41%)	116
Japanese	47 (32%)	147
Bikol	45 (28%)	162
Sochiapam Chinantec	43 (23%)	185
Mandarin	21 (18%)	115
English	27 (9%)	285

words, so function words can be used in recycling repairs as they are in English, to delay next item due. This leads to a high rate of recycling of bisyllabic function words, which, like function words in English, generally show initiation after recognizable completion. Consider Table 3.17, which shows how the relative frequency with which recycling repairs are initiated in bisyllabic function words and in bisyllabic content words in each language:

If we compare the number of bisyllabic function words and content words in which recycling repairs are initiated in the languages of our sample, we see that it is indeed the case that Indonesian speakers initiate repair far more often in bisyllabic function words than speakers of most of the other languages. Only Finnish has a higher percentage, and Finnish was the other language with an unusually high percentage of bisyllabic words in which recycling repair was initiated (see Table 3.16). This aspect of Finnish repair will be discussed further below.

We now consider the relative frequency of repair initiation in multisyllabic function and content words in the languages in our sample. If we look at Table 3.18 below, we see, first of all, that Mandarin and Sochiapam Chinantec speakers initiate repair in multisyllabic words far too infrequently for data from these languages to be relevant to this question. Of the other languages, Japanese, Bikol, and Finnish speakers are much more likely to initiate repair in multisyllabic content words than in multisyllabic function words. In Indonesian and English, on the other hand, speakers initiate repair in multisyllabic content words and function words with almost equal frequency.

Table 3.17 *Recycling repairs: Bisyllabic words by syntactic class*

	Function	Content	Total
Finnish	37 (79%)	10 (21%)	47
Indonesian	42 (64%)	24 (36%)	66
Bikol	18 (40%)	27 (60%)	45
English	10 (37%)	17 (63%)	27
Sochiapam Chinantec	14 (33%)	29 (67%)	43
Japanese	15 (32%)	32 (68%)	47
Mandarin	5 (24%)	16 (76%)	21

Just as with bisyllabic words so too with multisyllabic; Indonesian has an unusually high number of multisyllabic function words compared with most of the other languages in our sample. In fact, Indonesian by itself accounts for 31 percent of the multisyllabic function words in our data, although it only constitutes 11 percent of the multisyllabic words in which recycling repair is initiated. We can now see that the unusual finding that Indonesian speakers initiate recycling repairs after recognizable completion in longer words can be explained by the tendency to recycle function words to delay the next content word due.

Explaining English: Syntactic class and recognizable completion

So we are left with English as the language which cannot be fully predicted by our cross-linguistic findings about the correlation between word length and repair initiation. English speakers are unusually likely to initiate replacement repairs prior to recognizable completion, given that 62 percent of replacement repairs in English are of monosyllabic words, which would predict a tendency toward initiation after recognizable completion. We believe that English speakers are less attentive to length in replacements than they are to syntactic class; that is, in replacement repairs, English speakers far more frequently engage in initiation prior to recognizable completion whenever the word is a content word (adjective, noun or verb) and make use of initiation after recognizable completion whenever the word is a function word.

Table 3.18 Recycling repairs: Multisyllabic words by syntactic class

	Function	Content	Total
Japanese	6 (7%)	86 (93%)	92
Bikol	5 (7%)	62 (93%)	67
Finnish	9 (29%)	22 (71%)	31
Indonesian	11 (42%)	15 (58%)	26
English	4 (40%)	6 (60%)	10
Mandarin	0 (0%)	3 (100%)	3
Sochiapan Chinantec	0 (0%)	1 (100%)	1

Table 3.19 gives the relevant data for replacement repairs initiated in monosyllabic words.

The majority of replacement repairs initiated in monosyllabic function words are initiated after recognizable completion, and the majority of replacement repairs initiated in monosyllabic content words are initiated prior to recognizable completion. It seems likely that this difference relates to different interactional implications of function words and content words, where content words are more likely to be interactionally sensitive or inappropriate.

It would be desirable to compare English with other languages; however, for many of our languages the number of monosyllabic words in which replacement repairs are initiated is so small that the data cannot be considered representative, and any patterns they might seem to show must be treated with caution. Only Chinantec and Mandarin have more than five exemplars. Nonetheless, we have provided data for all five languages below in Table 3.20 (Indonesian has no replacements initiated in monosyllabic words).¹²

It appears that in Japanese, Finnish and Bikol, speakers show a tendency that is familiar to us by now: Monosyllabic function words are initiated after recognizable completion. Although the numbers are too small to be anything but suggestive, it looks as though speakers of these languages behave like speakers of English. Of course, unlike these three languages, the English corpus shows a large number of replacement repairs initiated in monosyllabic content words, a fact which makes English look, on the surface, quite different.

However, in both Sochiapan Chinantec and Mandarin, repairs in both function words and content words are generally initiated after recognizable completion. In other words, in neither language do speakers seem to be influenced by word type, which is

Table 3.19 English replacement repairs by word type for monosyllabic words

	Not recognizably complete	Recognizably complete	Total
Function words	3 (18%)	14 (82%)	17
Content words	14 (82%)	3 (18%)	17
Total	17 (50%)	17 (50%)	34

not surprising, given that we have seen that speakers of these two languages consistently favor initiation after recognizable completion. It is unclear to us at this point why English speakers (and probably speakers of other languages such as Japanese, Finnish and Bikol) seem to orient to syntactic class, and, thus, to interactional pressures, while Mandarin and Sochiapam Chinantec speakers orient mainly to word length. Further data from other languages may provide a clue as to why speakers of these two languages are so consistent in their preference for initiation at recognizable completion. Both are tone languages; one possible avenue to explore is whether tonality may be relevant.

A closer look at Bikol and Finnish: Word length and site of initiation

Bikol and Finnish both require further discussion because, although they match our predictions about site of repair initiation and word length, in Finnish recycling and both types of repair in Bikol speakers show no statistically significant preference for one or another site of initiation when all the data is grouped together. Because both of these languages have many multisyllabic words in their lexicons, according to the word-length criterion, we would have expected speakers to show a clear preference for initiation prior to recognizable completion. The answer to this puzzle lies in the distribution of word lengths of those words in which repair is actually initiated.

Table 3.20 *Replacement repairs by word type for monosyllabic words*

	Not recognizably complete		Recognizably complete		Total
	Function	Content	Function	Content	
Sochiapam	1	0	5	6	12
Chinantec					
Mandarin	0	0	9	18	27
Japanese	0	0	5	0	5
Finnish	0	0	2	2	4
Bikol	0	0	3	0	3

Table 3.21 *Word length in recycling repairs*

	Finnish		Bikol	
	Function	Content	Function	Content
Monosyllabic	38 (33%)	50 (30%)		
Bisyllabic	47 (41%)	45 (28%)		
Multisyllabic	31 (27%)	67 (41%)		
Total	116	162		

In both Finnish and Bikol, as shown in Table 3.21, the distribution of site of repair initiation according to word length is fairly even. Repairs initiated in monosyllabic words in these languages tend to be initiated after recognizable completion, those initiated in bisyllabic roots show fairly even division between the two types of initiation, and those in multisyllabic roots are mainly initiated prior to recognizable completion. As a result, although each syllable type follows our predictions, the data for the languages as a whole shows no preference for site of initiation of repair. Other languages in the sample did not show this type of balance between word lengths. In Japanese, for example, almost no repairs are initiated in monosyllabic words (5 percent), while in English (87 percent), Sochiapam Chinantec (76 percent) and Mandarin (79 percent) most repairs are initiated in monosyllables, and in Indonesian, as noted above, repair is most frequently initiated in bisyllabic words (56 percent).

Explaining monosyllabic repair initiation

One of the central findings of the current study is that speakers of languages of many different genetic, typological, and areal affiliations initiate repair after recognizable completion in monosyllabic words. This striking finding calls out for an explanation. In this section we proffer several hypotheses and examine evidence that might help us choose among them. In our discussion we focus primarily on recycling repairs.

Shortness vs beat of delay

As Schegloff (1979) has suggested, recycling is done, basically, to delay next item due. For whatever specific interactional or cognitive

reason a speaker recycles on a particular occasion, recycling always has the function of stopping progressivity of the turn-so-far and delaying next item due.

One obvious explanation, then, for our finding regarding site of initiation in monosyllabic words is that speakers shape their repair to give themselves a beat of delay. If we can say that one syllable gives roughly one conversational beat of delay, then initiating repair at the end of a syllable and repeating that syllable provides the speaker with one conversational beat of delay. Of course, speakers may also lengthen the syllable in which repair is initiated and may also allow silence to develop after the repair initiation, thereby increasing the temporal delay beyond a single beat; nonetheless, repeating a single syllable gives at least a minimum of one beat of delay.

This proposal finds support in research regarding the management of overlap. The majority of overlaps are resolved in one to two beats (Schegloff 2000c, 2002b),¹³ and one common device used to provide those one or two beats is turn-beginning recycling (Schegloff 1987c), in which the speaker recycles the first one or two syllables of the turn to secure a turn in the clear. This fact provides some support for our claim that many interactional difficulties can be resolved within one to two conversational beats and thus that recycling monosyllabic words may be used to provide the speaker with one beat of delay.

Moreover, in languages where function words precede content words, speakers have highly entrenched recurrent practices for delaying next content word due, practices which involve the recycling of high-frequency monosyllabic function words. For example, English speakers have as highly entrenched recycling devices "I-I," "the-the," "what-what," "that's-that's," "it's-it's" (but they do not have a comparable recurrent practice of, for example, "receive-receive"). These devices provide the speaker with one beat of delay just before a content word is due.

According to this explanation for our finding, it is beneficial for a speaker to initiate repair in such a way as to create one (or possibly two) extra conversational beats. In other words, a speaker may choose to initiate recycling repair after recognizable completion in a monosyllabic word, in order to achieve just this amount of delay. With longer words, recycling of only the first one or two

syllables of the word will produce the same kind of delay, thus allowing speakers to initiate repair earlier in multisyllabic words.

But a second possible explanation is also plausible. Monosyllabic words are short, often 200 milliseconds or less, and short words pose challenges to speakers who would like to initiate repair: Will the speaker decide to initiate repair in time to do so before the word reaches recognizable completion? If the speaker decides to initiate repair a bit late, it might not be possible to actually initiate the repair before the word is recognizably complete. The recycling still accomplishes delay, but the repair is initiated after recognizable completion because the speaker didn't recognize a need for delay until the last moment.

We thus have two clear possible explanations for our finding. We recognize that in any given case it is impossible to know when or why the speaker decided to initiate repair. However, when we look at the aggregate, we may be able to discover patterns which will provide clues as to what types of pressures speakers respond to most frequently. In the next section we explore three pieces of evidence that address this issue.

Three sources of evidence

The first piece of evidence comes from comparing sites of initiation for content words and function words in our languages. We have suggested that function words are most likely to be used to delay the next content word due and, therefore, to provide speakers with an extremely useful device. We would expect that content words would not be recycled with the same frequency as function words, because, while they may do important work in providing occasional delay for gaze organization, overlap resolution, etc., such functions are less frequent than the need to delay next content word due. In fact, in our sample of seven languages, in five of them speakers range from moderately to highly more likely to initiate repair in a function word than a content word. The two exceptions are Japanese, where function words follow content words and thus are less likely to be used to delay the next content word due, and Bikol, where one third of simple recycling repairs are initiated in verbs, and which has a rich system of verb prefixes which can be used to delay verbs.

A related piece of evidence comes from Indonesian. In Indonesian, most function words are bisyllabic. In spite of a cross-linguistic tendency for lack of preference in site of initiation in bisyllabic words, Indonesian speakers prefer initiation after recognizable completion for bisyllabic words, possibly because these tend to be function words that are used to delay the next content word due. This fact provides evidence that speakers indeed use function words to create a beat (or two) of delay.

English offers additional support for the idea that speakers use recycling to create a beat (or two) of delay. First, the vast majority of English recycling repairs (80 percent) are initiated in function words, which would be a surprising finding if there weren't some natural connection between function words and recycling, such as the one we have suggested of delaying next content word due. Second, even when content words are recycled, there is often evidence that they are being used to provide a beat of delay. Of the thirty-four monosyllabic content words in our collection, twenty-one were initiated after recognizable completion. Of those, the majority appear to be designed to provide delay in response to interactional pressures. Twelve are in overlap and attempt to bring the speaker into the clear. Another three occur in word searches within compound or complex noun phrases, where they are delaying the next noun due, and two are recyclings of the phrase "he goes," which delays an upcoming clause of reported speech. It therefore appears that English speakers often recycle complete words to provide a beat of delay, rather than because they have decided too late to initiate repair.

In addition, we have seen evidence from replacement repairs that English speakers are perfectly capable of initiating repair before recognizable completion in monosyllabic content words. One could argue that monosyllabic function words are even shorter than monosyllabic content words, due to frequency-induced phonological reduction, thus allowing for even less time to initiate repair; while this is no doubt true, the evidence from English still suggests that shortness alone is not the crucial factor in site of initiation in monosyllabic words.

We have thus found several facts to support the hypothesis that speakers initiate repair after recognizable completion in monosyllabic words in order to provide a beat of delay. While it is impossible to prove, or disprove, a claim that in some cases speakers

don't decide to initiate repair until they have reached the end of a word, and it is likely that this type of situation occurs, both when speakers initiate repair in a monosyllabic word and in a longer word, it is unlikely to be the explanation for the frequency with which we find repair initiation after recognizable completion in monosyllables.

However, Mandarin and Sochiapam Chinantec provide evidence that for some languages, neither of our explanations is appropriate. In these languages, repair is consistently initiated after recognizable completion in monosyllabic words, both content and function, and most repair is initiated in monosyllabic words. It seems that for these two languages speakers have a recurrent practice of initiating repair after recognizable completion, a practice that is only rarely overridden. We are not sure why this should be the case. Given that the two languages in question are both tonal, it may relate to tonality. Or it could be a kind of grammaticalization of a highly useful practice. Further research is needed to understand this issue more deeply.

Summary and conclusion

This is the first study of site of repair initiation from a cross-linguistic perspective that we know of. We began this study by noticing that there is a wide range of variation in relative frequency of site of initiation across the languages in our sample. We then asked the question: Is this variation just random variation, or do universal principles manifesting through language-specific means create the variation?

What we have proposed in this study is that there are indeed universal principles at work in shaping site of initiation patterns cross-linguistically. First, we noticed that different repair types can have very different patterns of initiation, probably relating to differences in interactional function.

Second, we noticed that length of word is one factor that correlates with site of initiation. We found that in both recycling and replacement repairs monosyllabic words tend to be repaired after recognizable completion in all the languages of our study; furthermore, multisyllabic words tend to be repaired prior to recognizable completion in all but one of the languages that have

enough multisyllabic words to show a pattern. Indonesian is the only exception; in this language, speakers show no preference for initiation before recognizable completion in multisyllabic words (see Table 3.11). We also observed that in most languages in our sample speakers show no preference for site of initiation in bisyllabic words. In Indonesian, however, which has an unusually high frequency of repair initiated in bisyllabic words, speakers strongly prefer initiation after recognizable completion.

This means that in languages in which speakers initiate repair mainly in monosyllabic words (such as Sochiapam Chinantec, Mandarin, and English), they tend toward initiation after recognizable completion, while in languages in which speakers prefer initiation in multisyllabic words (such as Japanese), they tend toward initiation prior to recognizable completion. On the other hand, if speakers of a language shows no preference as to the length of the words in which they initiate repair (such as in Bikol and Finnish), they will also show no overall preference for a single site of repair initiation, even though for each of the three categories of word length (monosyllabic, bisyllabic, and multisyllabic) they follow the patterns we have observed.

Third, we noticed that replacement repair tends to be initiated prior to recognizable completion in four of our seven languages (English, Indonesian, Japanese, and Finnish). Sochiapam Chinantec and Mandarin, in which speakers almost always initiate repair after recognizable completion, do not participate in this pattern, nor does Bikol, where speakers show no preference.

We offered several possible explanations for these cross-linguistic patterns. First, we suggested that speakers tend to initiate repair after recognizable completion in monosyllabic words because of late decisions to initiate repair. Since speakers may not decide to initiate repair until they have gotten partway through a word, if those words are monosyllabic, then by the time the speaker chooses to initiate repair, the word is recognizably complete. It is worth noting that high-frequency function words are often somewhat and sometimes radically phonologically reduced, at least in English (Juratsky *et al.* 1998). Phonological reduction might make it more difficult to initiate repair prior to recognizable completion.

Second, we explored the possibility that speakers initiate repair after recognizable completion in monosyllabic words because they want to create a beat of delay. We suggested that there may be a tendency for monosyllabic words to be function words, and that

function words tend to be recycled to delay next content word due; in such cases – in languages for which function words precede content words – speakers may want a beat (or two) of delay, since one of the main purposes of recycling in these instances is to provide a temporal delay. This same reasoning could explain why in Indonesian repair tends to be initiated in bisyllabic words after recognizable completion, since in Indonesian function words are frequently bisyllabic. Clearly such reasoning does not apply to Japanese, since in Japanese function words follow content words. Japanese has developed other techniques for accomplishing delay (Fox *et al.* 1996).

We proposed that while either of these explanations could be the relevant one in any given case, the overall patterns of repair initiation that we have observed suggest that speakers frequently recycle monosyllabic words to achieve a beat of delay, and that repair initiation after recognizable completion is a recurrent practice. However, Mandarin and Sochiapam Chinantec, where speakers have a consistent practice of initiating repair after recognizable completion in all word-lengths, under almost all circumstances, seemed to suggest the possibility that in some languages speakers may have generalized the pattern to be a fixed practice.

We also suggested that multisyllabic words provide many opportunities to initiate repair before coming to the recognizable completion of the word. It may be that in most cases of repair initiation, the speaker needs only one or two extra conversational beats, and with multisyllabic words a speaker doesn't need to produce the whole word to get those beats. We further suggested that speakers favor early initiation in replacement repairs in several languages just because replacements are generally of content words, which are generally longer in most languages in our study. Additionally, speakers may choose to initiate repair sooner so as not to be held accountable for alternate formulations or so as not to produce inappropriate terms. In Mandarin and Sochiapam Chinantec this tendency is not expressed.

We hope to have demonstrated that site of repair initiation can be a window onto differences across languages, as well as onto universal principles of repair organization. We also hope to have provided evidence that sometimes differences across languages mask the universal interactional patterns that organize them, and that research on cross-linguistic differences also benefits from a search for those universal patterns (cf. Sidnell 2007c).

Acknowledgments

We would like to thank Vincent Sarich for statistical and editorial assistance, Emmanuel A. Schegloff for clarifying comments and for providing the original inspiration for this study, and Jack Sidnell for inviting us to participate in this volume and for insightful comments on an earlier draft. Funding was provided by NSF grant number BCS0406512.

Authorship

Barbara Fox analyzed the English data. Fay Wouk analyzed the Indonesian data. Makoro Hayashi analyzed the Japanese data. Steven Fincke analyzed the Bikol data. Liang Tao analyzed the Mandarin data. Maria-Leena Sorjonen and Minna Laakso analyzed the Finnish data. Wilfrido Flores Hernandez analyzed the Sochiapam Chinantec data. Barbara Fox and Fay Wouk wrote the majority of the paper with data and input provided by the other co-authors.

Notes

- 1 Replacements and deletions differ in that in replacements the word type remains although the token has been replaced, while in deletions the word type is eliminated.
- 2 This has been amply demonstrated for English by Jasperson (1998).
- 3 This might seem to be a substantial change from Schegloff's proposal, since the category post-beginning has been expanded from including just the first sound of the word to including all but the very last sound. However, in actual fact, at least for English the difference is not as great as it sounds. It turns out that for English (the language on which Schegloff based his original proposal), initiation after the first syllable but prior to recognizable completion is an extremely rare occurrence; in our corpus of 339 simple repairs, only eight were initiated in these in-between locations. The same is true of Sochiapam Chinantec and Mandarin. It's only when we turn to languages with large numbers of multisyllabic words that the middle category becomes robust.
- 4 By inappropriate we do not mean offensive to the recipient; we simply mean that it is not the word that the speaker finds best suited to the context.
- 5 And in some cases the speaker may wish to be accountable for an inappropriate or offensive term.
- 6 For Japanese, the relevant category is actually the *mora*, not the syllable. All Japanese counts in the following discussion represent *mora*

Initiation in same-turn self-repair

rather than syllable counts. However, for simplicity of exposition, we will use the term "syllable" throughout when referring to all languages in the study. The term *mora* will only be used in discussions restricted to Japanese.

- 7 All three of these languages have a range of word lengths; 80 percent of words in the modern Mandarin lexicon are compound words, thus usually bisyllabic; Sochiapam Chinantec has mainly monosyllabic roots but makes use of a wide range of prefixes to form bisyllabic words; the English lexicon clearly has many words of two or more syllables. However, according to our findings, in these languages repair is initiated primarily in monosyllabic words. Studies of the frequency of different word lengths in everyday conversation in these languages might shed light on this discrepancy.
- 8 These differences are no doubt in part a reflection of the average length of words in general the different languages.
- 9 Word length in Mandarin was determined as follows. If the syllable (where repair is initiated) and its following syllable(s) form a tight unit in that they cannot be further divided into sub-components (Adj-N), then it is a bisyllabic or multisyllabic word. If the syllable and its following syllable(s) form a relatively loose unit that allows insertion of grammatical elements into it (i.e., the site of initiation is a syllable that has a meaning of its own to contribute to the whole utterance), then this syllable is a monosyllabic word.
- 10 Because content words are open class, there are a larger number of potential candidates in any given context than there are for function words. In addition, content words are generally of lower frequency than are function words. Thus speakers face a greater challenge in selecting the appropriate term.
- 11 In fact, it is more puzzling that Sochiapam Chinantec speakers do not show this preference than that Mandarin speakers do, given the many similarities we have seen between these two languages where repair is usually initiated in monosyllabic words. It would appear that the difference relates to the presence of verbal prefixes in Sochiapam Chinantec and the frequency with which verbs are broken off after the prefix but before the root as a way of delaying the next verb due. However, a complete discussion of this matter is beyond the scope of this paper.
- 12 Although Mandarin has thirty simple replacement repairs, three had to be excluded from this analysis, as the syntactic class of the word that was replaced was ambiguous.
- 13 Resolution in the first two beats is done without recourse to techniques of competitive production, while those that go over two beats typically involve competitive production (Schegloff 2000c, 2002b).