A coding scheme for other-initiated repair across languages

Mark Dingemanse*, Kobin H. Kendrick, N. J. Enfield

Abstract: We provide an annotated coding scheme for other-initiated repair, along with guidelines for building collections and aggregating cases based on interactionally relevant similarities and differences. The questions and categories of the scheme are grounded in inductive observations of conversational data and connected to a rich body of work on other-initiated repair in conversation analysis. The scheme is developed and tested in a 12-language comparative project and can serve as a stepping stone for future work on other-initiated repair and the systematic comparative study of conversational structures.

Keywords: interactional linguistics, conversation analysis, other-initiated repair, typology, coding

1 Introduction

Other-initiated repair is one of the key mechanisms that people use for the real-time resolution of problems of speaking, hearing and understanding in conversation. The basic structural facts of other-initiated repair are well-understood, making it an excellent locus for systematic comparative research (Schegloff 1993). This coding scheme, the introduction, and the other papers in this special issue demonstrate one approach to the systematic comparison of other-initiated repair across languages.

The conceptual framework of the project is detailed in the introduction to the special issue. Practices of other-initiated repair are realised with language-specific resources (Sidnell 2007). While the coding scheme we present here is informed by existing English-based taxonomies of formats for other-initiation of repair (as found in Schegloff, Jefferson, and Sacks 1977; Purver 2002; Sidnell 2010, among others), it is designed to enable researchers to distinguish language-specific practices from properties that apply across languages. To aid cross-linguistic applicability, the scheme captures the constitutive properties of practices of other-initiated repair, in the form of linguistic and sequential attributes that qualitative analysis has shown to be reliably identifiable across languages.

The scheme is one outcome of a 12-language comparative project on other-initiated repair (Dingemanse et al. 2015). At the base of it lies an iterative process of qualitative analysis, data sessions, and coding design. The scheme’s questions and analytical categories are grounded in inductive observations of conversational data and in prior work on other-initiated repair. To show this, the scheme is richly annotated with examples from published transcripts and with explanations and references to prior research where relevant. We also include brief instructions for sampling interactions and building collections, as well as some
basic conventions for the application of the coding scheme. Tried and tested against a broad and diverse sample of social interactions, this coding scheme can be used to test existing proposals and formulate new research questions.

Distributional evidence has always played a key role in conversation analysis (Schegloff, Jefferson, and Sacks 1977; Sacks and Schegloff 1979, inter alia), and formal coding is one way to make explicit the activity of creating collections and aggregating cases based on interactionally relevant similarities and differences. As formal coding of interaction is increasingly common in conversation analysis (Heritage 1999; Stivers and Enfield 2010; Stivers 2015), published coding schemes serve a dual purpose: (i) they document which analytical choices have been made and how the coding categories have been operationalised, ensuring evaluability; (ii) they can serve as a reference point and a stepping stone for future work in the domain, enabling cumulative progress.

2 Guidelines for collecting and coding sequences of other-initiated repair

We define the core of other-initiated repair in language-neutral terms as follows: “a sequence in which a turn T0 signals some trouble in a prior turn T-1 and is treated as making relevant the provision or ratification of a repair solution in a next turn T+1” (Dingemanse and Enfield 2015:99). The terms T-1 and T+1 are defined relative to T0: T-1 is the trouble source turn targeted by T0, often but not always the turn directly preceding it (there may be intervening material). Likewise, T+1 is the turn providing a candidate repair solution in response to T0, often but not always the directly following it.

Throughout the coding scheme, we refer to the participants in the repair sequence as “A” (the party producing T-1 and typically also the self-repair at T+1) and “B” (the party initiating repair at T0). Extract 1 illustrates the elements of a repair sequence using a classic example from Schegloff, Jefferson and Sacks (1977:367):

Extract 1 T0:27

1 A: Oh Sibbie’s sistuh hadda ba:by bo:way. T-1
2 B: Who? T0
3 A: Sibbie’s sister. T+1
4 B: Oh really?

The coding of conversational data should proceed from a deep qualitative understanding of the phenomenon under investigation, and should be based on a systematic collection of cases identified in a clearly specified sample of talk-in-interaction. To minimise sampling bias, collections should be based on data from a range of different speakers in different interactions. Here are three general guidelines for the creation of collections for comparative purposes:

1. Cases should come from continuous segments of interaction. For instance, if you are working with a 10-minute segment, you must include all examples in that stretch. In the context of a comparative project, this is especially important if some estimate of relative frequency is needed, for instance to decide how much data to sample.
2. Cases should be collected liberally. If you are unsure whether something is a case you should note it down anyway — it is easier to systematically exclude cases later on, while it is a lot of work to go back and look for cases that you failed to include (Schegloff 1996; 1997; and see Dingemanse and Enfield 2015; Kendrick 2015a for how the project documented here dealt with inclusion and exclusion).
3. Cases should be collected from equally-sized segments taken from as many different interactions as possible (including both dyadic and multi-person, quiet and noisy, with and without parallel activities, etc.), so that sampling bias is minimised and coverage is broad.
Following these guidelines will result in a collection that is maximally representative of the diversity of the conversational corpus as a whole, and maximally suited for cross-linguistic comparison. Additionally, three coding conventions help ensure the integrity of data and analysis:

1. Coding should be done on the basis of a coherent and systematic collection of candidate sequences of other-initiated repair (see guidelines above).
2. Coding should be based on the fullest representation of the data available. Although the coding scheme includes a simplified transcript of the trouble source turn and repair sequence for the purposes of quality control, for most of the questions it is necessary to rely on the original transcripts, audio, and video.
3. Coding should be done conservatively. When in doubt, follow the data session principle: for any coding decision, you should be able to defend your choice based on the evidence available within the interaction. If that is not the case, code as “can’t tell”. If a question is not applicable, code as “NA”.

Following these coding conventions promotes the systematicity, comprehensiveness and clarity of the analysis. The data session principle is particularly important as it helps establish a common standard of rigor and reminds analysts that aggregating and classifying cases is always rooted in a firm qualitative understanding.

3 The coding scheme

Part A: basic data

This section records the basic data for every single repair initiation and its trouble source turn and repair solution. Unique identifiers are used to locate cases and refer to them in analysis or data sessions. Transcripts and translations help make cases intelligible to analysts working on other languages.

A0. Language
A1. Unique identifier
A2. T-1 transcript (also include relevant bodily behaviour)
A3. T-1 translation
A4. T0 transcript (also include relevant bodily behaviour)
A5. T0 translation
A6. T+1 transcript (also include relevant bodily behaviour)
A7. T+1 translation

Part B: T0 (repair initiation)

B1. What is T0?

OIR English Extract 1 (Kendrick 2015a:166)¹
QNR Dutch Extract 2 (Dingemanse and Enfield 2015:100)
other Siwu Extract 24 (Dingemanse 2015:253)

The function of this question is to distinguish core cases of other-initiated repair (OIR) from question-formatted news receipts (QNR, like really?, he did?) and from repair-initiator formats used for other purposes (e.g. huh? as a response to summons (Schegloff 1997), WHAT! as an exclamation of surprise). See Kendrick’s (2015a) article on English for one possible articulation of the boundary between core and peripheral cases of OIR.

¹ These references point readers to published examples of each value for each coding question. Virtually all of the examples can be found within this special issue, which is fully and freely available online through Open Linguistics.
[Only answer B2 if B1 = “OIR” or “other”]

B2. What other action does T0 perform?

- no other action: Lao Extract 1 (Enfield 2015:120)
- surprise/disbelief: Yélî Dnye Extract 12 (Levinson 2015:397)
- disaligning action: Italian Extract 15 (Rossi 2015:277)
- non-serious action: English Extract 23 (Kendrick 2015a:183)
- other

Other-initiations of repair can perform ancillary actions: they may perform a display of surprise or disbelief (Selting 1987; Wilkinson and Kitzinger 2006), disaligning actions, such as disagreements, challenges, rejections, or preliminaries thereto (Schegloff, Jefferson, and Sacks 1977; Drew 1995; Schegloff 1997; 2007), non-serious actions like jokes, puns, or teases (Kendrick 2015a; Gisladottir 2015), or other actions.

[All remaining questions are only answered for core OIR cases, i.e. B1 = “OIR”]

B3. What is the position of T0 in the in-progress OIR sequence?

- one and only: Russian Extract 1 (Baranova 2015:557)
- first of non-minimal: Siwu Extract 2 line 3 (Dingemanse 2015:235)
- other: English Extract 2 lines 7, 13, 17 (Kendrick 2015a:167)
- last of non-minimal: Yélî Dnye Extract 2 line 4 (Levinson 2015:389)

An other-initiation of repair may occur on its own or as part of a more extended (non-minimal) repair sequence, in which case it may be the first, the n-th (other), or the last in the extended sequence (Enfield, Drew, and Baranova forthcoming).

B4. How does T0 target the problem to be repaired?

- open: Murrinh Patha Extract 10 (Blythe 2015:297)
- restricted: Italian Extract 8 (Rossi 2015:267)
- alternative question: LSA Extract 20 (Manrique 2015:747)
- external: Italian Extract 9 (Dingemanse and Enfield 2015:100)

Open repair initiations signal a problem while leaving open what or where it is, while restricted repair initiations further specify the nature or location of the problem (Drew 1997; Benjamin 2013:19–66). Alternative questions are closely akin to restricted repair initiations and invite a selection from among alternatives (Koshik 2005). External RIs target things that were not expressed in the trouble-source turn (Schegloff, Jefferson, and Sacks 1977:369n15). Restricted repair initiations include category-specific question words, full repeats, partial repeats and candidate understandings (Schegloff 1993) —subtypes that are distinguished through subsequent questions (B6-B8, D2-D4).

[Only answer B5 if B4 = “open”]

B5. What type of open repair initiation is this?

- interjection: Cha’palaa Extract 4 (Floyd 2015:473)
- question word: Icelandic Extract 5 (Gisladottir 2015:316)
- formulaic: Russian Extract 12 (Baranova 2015:564)
- visible only: LSA Extract 1 (Manrique 2015:727)
An open repair initiation may be an interjection like *huh?* or a closed-mouth *mm?* (Dingemanse 2015:238–9), or it may include a question word like *what?* (Schegloff, Jefferson, and Sacks 1977). It may also be a formulaic expression like *sorry?* in English (Robinson 2006) or *prostitute* in Russian (Baranova 2015), or a non-verbal visible sign like a head movement or facial expression (Seo and Koshik 2010; see also Manrique 2015). Visible actions that accompany verbal forms are treated in questions E3 and E4.

[Only answer B6-B8 if B4 = “restricted”, “alternative question”, or “external”]

B6. Does T0 include any repeated material from T-1?

- **full** Lao Extract 20 (Enfield 2015:132)
- **partial** Icelandic Extract 12 (Gisladottir 2015:320)
- **no** Cha’palaa Extract 15 (Floyd 2015:481)

Repetition may serve different functions in repair initiation (Jefferson 1972; Wu 2009; Robinson and Kevoe-Feldman 2010; Robinson 2013). A full repeat includes all material from T-1 while a partial repeat includes only some material from T-1. If there are deictic shifts (e.g., *you* → *I*) this still counts as repetition.

B7. Does T0 involve a content-question word?

- **yes** Murrinh Patha Extract 7 (Blythe 2015:292)
- **no**

The content-question word can stand alone as a lexical TCU (e.g., *who?*) or it can occur as an argument in a clause (e.g., *who’s that?*). Content-question words are also known as category-specific question words (Schegloff 1993). Content-questions words that serve as open other-initiations of repair (see B5) are not relevant to this question.

B8. Does T0 make (dis)confirmation relevant?

- **yes** Cha’palaa Extract 11 (Floyd 2015:480)
- **no** English Extract 13 (Kendrick 2015a:173)

Restricted repair initiations that offer a solution for confirmation (Schegloff 2004) include candidate understandings as well as full or partial repeats (or “candidate hearings”, Dingemanse, Blythe, and Dirksmeyer 2014). Whether confirmation or disconfirmation indeed follows is tracked separately in D4.

B9. Is there ‘added’ explicit marking of the T0 function?

- **yes** Lao Extract 32 (Enfield 2015:139)
- **no**

Some repair initiations include linguistic material that is closely connected to (specialised functions of) repair initiation. For instance, candidate understandings may be prefaced with *y’mean* in English (Schegloff et al. 1977; Benjamin 2012), though this is rare (Kendrick 2015a), subject and object particles in Korean may mark upcoming disagreement or surprise (Kim and Kim 2014), and final particles in Mandarin Chinese (Wu 2006) and Lao (Enfield 2015) may mark interactional stances.

---

2 For binary questions like this we usually provide examples for the most informative answer only (here, ‘no’ could be trivially illustrated by virtually any other format for repair-initiation). We do provide an example for the “no” answer if it exemplifies a borderline case that helps the reader understand some of the tricky analytical decisions we have had to make.
B10. Is T0 simple or complex?

<table>
<thead>
<tr>
<th>Simple</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siwu Extract 20 (Dingemanse 2015:249)</td>
<td></td>
</tr>
</tbody>
</table>

The turn in which repair is initiated may be simple (a single repair initiation occupying a single TCU), or complex, involving combinations of formats (“doubles”, Kim 1999). To qualify as complex, a repair initiation should occupy one intonation unit.

**Part C: T-1 (repairable or trouble source turn)**

C1. What is the sequential status of T-1?

<table>
<thead>
<tr>
<th>FPP</th>
<th>Russian Extract 7 (Baranova 2015:562)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPP</td>
<td>Italian Extract 4 (Rossi 2015:263)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

The sequential environment of repair initiations may bear on why and how repair is other-initiated (Drew 1997). The trouble source turn may be a question or another type of first pair part (FPP), or the trouble source turn may be a response to a question or another type of second pair part (SPP) (see Schegloff 2007:100–106,149–155 on the initiation of repair at these positions).

C2. Is there a cut-off, unit restart, and/or um/uh in T-1?

<table>
<thead>
<tr>
<th>Yes</th>
<th>Italian Extract 1 (Rossi 2015:258)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Can’t tell</td>
<td></td>
</tr>
</tbody>
</table>

Self-repair may indicate troubles of speaking and may bear on when and how repair is initiated (Schegloff 1992:1302 where an other-initiation of repair locates a self-repair as a trouble source).

C3. Is there noise or overlapping talk during T-1?

<table>
<thead>
<tr>
<th>Noise</th>
<th>Cha’palaa Extract 1 (Floyd 2015:469)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap</td>
<td>Siwu Extract 3 (Dingemanse 2015:237)</td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Can’t tell</td>
<td></td>
</tr>
</tbody>
</table>

Noise or overlapping talk may cause troubles of perception and may bear on when and how repair is initiated (Schegloff 2000). Something should only count as “noise” or “overlap” if the analyst is able to argue that it could disturb the audibility of T-1 for participant B.

C4. Could the trouble be due to not being able to see something, for example, a speaker or referent?

<table>
<thead>
<tr>
<th>Yes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Can’t tell</td>
<td></td>
</tr>
</tbody>
</table>

This question, closely akin to C3, is designed to capture troubles of perception in the visual modality, which may bear on when and how repair is initiated. It was hypothesized to be especially applicable in
signed languages. Subsequent analysis reveals that in Argentinian Sign Language, such problems do indeed occasionally occur (as when the line of sight between A and B is blocked), but tend to be resolved most commonly not by B other-initiating repair but by A self-repairing (Manrique 2015).

C5. If yes, briefly describe the visual problem from C4

[fulltext]

Examples are “A and B are at considerable distance from each other”, “the referent A talks about is not in B’s visual field”.

C6. During T-1, is B involved in a parallel course of action that is demanding on B’s attention?

yes Italian Extract 3 (Rossi 2015:261)
no
can’t tell

A parallel course of action such as preparing food, attending to a computer, or being engaged in a conversation with someone other than A may be a potential cause of troubles in hearing or understanding and may bear on when an how repair is initiated (Rossi 2015:261; Blythe 2015:297–299).

C7. If yes, briefly describe the parallel course of action that B is involved in

[fulltext]

Examples are “B is attending to a computer”, “B is engaged in another conversation”.

Part D: T+1 (repair solution)

D1. Does T+1 repeat any material from T-1?

full Yéli Dnye Extract 1 (Levinson 2015:388)
partial Russian Extract 4 (Baranova 2015:560)
no

Repair solutions regularly feature self-repetition of material from the trouble source turn (Curl 2005). A full repetition includes all material from T-1 while a partial repeat includes only some material from T-1. If there are omissions of dispensable items (Schegloff 2004) this still counts as repetition (and this omission is tracked in D2).

[Only answer D2 if D1 = “full” or “partial”]

D2. If a repeat, have dispensable items been left off?

yes Siwu Extract 1 (Dingemanse 2015:234)
no Lao Extract 8 (Enfield 2015:125)

Some items—including vocatives, turn-initial but, well and oh, and turn-final tags—may become ‘dispensable’ on a second saying (Schegloff 2004; Dingemanse 2015:251, Table 7).
D3. Is T+1 a modified version of T-1?

- **yes** Icelandic Extract 3 (Gisladottir 2015:314)
- **no** Yélî Dnye Extract 1 (Levinson 2015:388)

Repair solutions may involve not just repetition of a trouble source but also modification and specification (Mazeland and Zaman-Zadeh 2004). T+1 can be a version of T-1 that is modified (i.e., part or all of T-1 was replaced or rephrased, or something was added) or not modified (i.e., a verbatim repetition).

D4. Does T+1 include a (dis)confirmation of T0?

- **yes** Murrinh-Patha Extract 9 line 8 (Blythe 2015:296)
- **no**
- **can’t tell**

Some repair initiations offer a solution for confirmation or ‘checking’ (Schegloff, Jefferson, and Sacks 1977; Schegloff 2004). The format of the repair initiation is coded in B8; here we track whether (dis)confirmation indeed follows. Code conservatively: “yes” only if T+1 includes a ‘yes/no/indeed’ type item, a head nod/shake, or a repetition (+/- negation); “no” only if T+1 does no confirmation at all or if it does indirect (dis)confirmation, e.g. by offering an alternative.

D5. Is there ‘added’ explicit marking of the T+1 function?

- **yes** Lao Extract 37 (Enfield 2015:141)
- **no**

Repair solutions may include additional linguistic material that is not part of the redoing of T-1 itself, but performs an additional function such as framing the resayings or indicating the speaker’s stance. See Enfield’s (2015:141–2) discussion of final particles that mark T+1 in Lao.

**Part E: Visible behaviour**

E1. During T-1, is A gazing at B at some point?

- **yes** Italian Extract 3 (Rossi 2015:261)
- **no**
- **can’t tell**

Gaze behaviour may bear on how and why repair is initiated (Egbert 1996).

E2. During T-1, is B gazing at A at some point?

- **yes**
- **no** Italian Extract 3 (Rossi 2015:261)
- **can’t tell**

Gaze behaviour may bear on how and why repair is initiated (Egbert 1996). The gaze behaviour of B, and the alignment and attention of the participants that this indexes, is one factor that appears to influence the selection between interjection and question-word open other-initiations of repair (Dingemanse, Blythe, and Dirksmeyer 2014:16–7; Blythe 2015:302–305; Kendrick 2015a:180n8).
E3. Around T0, does B move their body or head so as to increase perceptual access?

<table>
<thead>
<tr>
<th></th>
<th>English (Figure 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td></td>
</tr>
<tr>
<td>can't tell</td>
<td></td>
</tr>
</tbody>
</table>

Other-initiation of repair may come with movements of the body or head, such as leaning forward or turning towards A (Seo and Koshik 2010; Rasmussen 2014; Li 2014; Rossi 2015:262; Kendrick 2015a:178). These movements may start earlier or end later than the verbal component of the repair initiator, and they may increase perceptual access to A's anticipated redoing of T-1 as well as index (renewed) attention on B's part.

![Figure 1](image1.jpg)

*Figure 1.* The pairs of images above show speaker B's head position during T-1 and immediately after T0. The first pair illustrates a forward head extension and would be coded as “yes” while the second pair illustrates no relevant movement of the head or body and would therefore be coded as “no”.

E4. Around T0, is there easily noticeable facial action by B?

<table>
<thead>
<tr>
<th></th>
<th>Lao (Figure 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td></td>
</tr>
<tr>
<td>can't tell</td>
<td></td>
</tr>
</tbody>
</table>

Other-initiation of repair may come with facial displays such as a raising or furrowing of the eyebrows, squinting, or nose wrinkles (Ekman 1979; Wiener et al. 1972). These facial displays may start earlier or end later than the verbal component of the repair initiator.

![Figure 2](image2.jpg)

*Figure 2.* The speaker's facial expression during two different repair initiating turns (T0). The left panel illustrates a raising of the eyebrows during a repair initiator *haa2* ‘huh?’ and would be coded as “yes”. The right panel illustrates no easily noticeable facial action and would be coded as “no”.


E5. Is any behaviour held by B from around T0 at least until the beginning of the repair solution (eye gaze, head position, body posture, manual co-speech gesture or signs, facial articulation)?

yes Examples of this phenomenon can be found in Floyd et al. (2015).
no can't tell

Movements of the eyes, heads, hands or body may be temporarily and meaningfully held static while the repair sequence plays out (Floyd et al. 2015). This may be until the repair solution T+1 or even until the solution is ratified in a next turn T+2 (usually knowledge uptake or resumption of progressivity).

Part F: Sequence

F1. What is the participation framework during this sequence?

   dyadic LSA Extract 6 (Manrique 2015:743)
   multi-person Lao Extract 3 (Enfield 2015:123)
   can't tell

Other-initiated repair, like turn-taking, may be sensitive to the number of parties involved (Sacks, Schegloff, and Jefferson 1974; Egbert 1997). Coding should be done conservatively: if more than two people can be heard or seen, the choice would be multi-person, unless there is good evidence that the sequence is really just dyadic.

F2. Do more than two people contribute to this OIR sequence?

   yes Icelandic Extract 12 (Gisladottir 2015:320)
   no

Sometimes a repair solution may be provided by a party other than the original speaker (Egbert 1997; Bolden 2011), in which case more than two people contribute to the sequence.

F3. Is there intervening material between T-1 and T0? (This could be a TCU, a full turn or even multiple turns.)

   yes English Extract 15 (Kendrick 2015a:174)
   no

Sometimes material (a TCU, a full turn, or multiple turns) intervenes between a trouble source turn or TCU and a repair initiation (Schegloff 2000; Wong 2000). In English, this has been shown to occur in close to a quarter of all cases (Kendrick 2015b:5) and to affect the design of the repair initiation (Benjamin 2012; Kendrick 2015a:174).

Acknowledgements: This coding scheme was developed within the framework of the research program Human Sociality and Systems of Language Use, led by Nick Enfield at the Max Planck Institute for Psycholinguistics. It has benefited from the sustained input of all project participants: Julija Baranova, Joe Blythe, Mark Dingemanse, Tyko Dirksmeyer, Paul Drew, Nick Enfield, Simeon Floyd, Rósa Gísladóttir, Robin Kendrick, Steve Levinson, Elizabeth Manrique and Giovanni Rossi. We thank two anonymous reviewers for their helpful comments. We also thank Jeff Robinson for useful feedback. The research was funded by the European Research Council (projects HSSLU 260583 and INTERACT 269484) and supported by Steve Levinson’s Language and Cognition Department at the Max Planck Institute for Psycholinguistics.
References


