

Chapter 8

Potential pitfalls of interpreting data from English-French tandem conversations

Sylwia Scheuer

University of Paris 3 – Sorbonne Nouvelle

Céline Horgues

University of Paris 3 – Sorbonne Nouvelle

The chapter focuses on methodological issues involved in analysing, coding and interpreting data from the *Spécificités des Interactions verbales dans le cadre de Tandems linguistiques Anglais-Français (Characteristics of English/French spoken tandem interactions)* corpus of English-French tandem exchanges. Each of the 21 tandem pairs recorded consisted of a native speaker of English and a native speaker of French. The participants were video and audio recorded while performing tasks (conversation and reading) in both languages. So far, two major threads of research on the corpus data have emerged: corrective feedback and communication breakdowns. We have attempted to gain insights as to when or why corrective feedback is given to the non-native tandem partner and when or why communication between the partners gets compromised. Findings from those previous thematic areas serve as the basis for the present study. The major challenge we have encountered in conducting the analyses is the ambiguity and complexity of our conversational data. Both corrective feedback and communication breakdowns may have multiple – and not always obvious – causes and may or may not be clearly signalled by the participants. In the chapter, we discuss the various problems we faced and addressed while coding the data, as well as how the methodological choices we made affect our results and conclusions. The discussion is amply illustrated with examples from the corpus.

Keywords: Tandem learning, corrective feedback, communication breakdowns, data coding, NS-NNS interactions



1 Introduction

Tandem learning is “an arrangement in which two native speakers of different languages communicate regularly with one another, each with the purpose of learning the other’s language” (O’Rourke 2005: 434). Consequently, tandem interactions constitute a unique collaborative language-learning environment, which is based neither on the socially institutionalised teacher-learner hierarchy nor on the exact symmetry of peer interactions, where learners share their first language (L1) and their target second language (L2). Instead, it is based on role-reversibility and solidarity between the two tandem partners, each of whom will construct two roles throughout the conversation exchange and, more generally, throughout their tandem history: the role of the (relative) expert when speaking in their mother tongue and the role of the learner, or the less proficient speaker, when speaking in the L2. The fact that each participant gets to wear the hat of both the native speaker (NS) and the non-native speaker (NNS) at some point in the interaction makes their relationship essentially non-hierarchical. Language-expertise asymmetry is only contextual (the conversation invariably switches from one’s L1 to L2, or the other way round, within a short period of time), which makes tandem exchanges also different from the classic NS-NNS conversational setting, where the expert-novice relationship is not reversible.

The database that the present contribution draws on to discuss such interactions is the *Spécificités des Interactions verbales dans le cadre de Tandems linguistiques Anglais-Français* (SITAF: *Characteristics of English/French spoken tandem interactions*) corpus, in which we collected linguistic material – both video and audio recorded – from face-to-face conversational exchanges held by 21 pairs of undergraduate students at the University of Paris 3 – Sorbonne Nouvelle. Each such tandem consisted of a NS of English and a NS of French. By virtue of containing largely unscripted L1-L2 productions, the corpus offers ample opportunities for various types of analyses of NS-NNS interactions, including studies of corrective feedback (CF) and communication breakdowns (CBs). It is those two, overlapping, research areas that the chapter focuses on, with a view to presenting various methodological challenges that researchers can face when coding and interpreting data.

We equate corrective feedback with the verbal provision of negative evidence. Negative evidence, in turn, is defined as “the type of information that is provided to learners concerning the incorrectness of an utterance” (Gass 2003: 225) – in other words, information as to what is not possible, or not deemed acceptable, in a given language. This can be illustrated with the following example from the

8 Potential pitfalls of interpreting data from E-F tandem conversations

SITAF corpus, where an American participant comments on his French partner's renditions of the 'th' sounds:

- (1) NS: The only suggestion that I could make for you was the /θ/ sound [...] I could completely understand you, and everyone else could, but... erm... instead of ['zi] it's ['ði:].

Here, the L2 English learner gets corrected on a pronunciation issue which, by the NS's own admission, did not cause any communicative turbulence.

By definition, communication breakdowns do hamper communication, at least at some point in the conversation. Our conception of CBs includes all cases where the listener has difficulty or is incapable of grasping the meaning of an utterance as seemingly intended by the speaker, and makes that difficulty somehow visible or audible. Naturally, what the speaker truly means can be a matter of speculation, although the study of the broader context in which the interaction takes place usually sheds sufficient light on the matter. The following exchange serves as an example of a successfully resolved communication breakdown in our corpus:

- (2) NNS: You know, when people are calling you you are sometimes hungry [* ['hʌŋgri]].
NS: [laughing] Wait, the person you are calling is hungry?
NNS: No, no, no, no, the person which is called.
NS: Oh, is it angry [hyperarticulation]?
NNS: Yeah! Sorry, sorry for my accent.

As example (2) demonstrates, communication breakdowns arising from the speech of a NNS will, to a large extent, also involve corrective feedback. Very often, a CB instance will actually trigger an input-providing corrective sequence, such as the hyperarticulation of the mispronounced adjective above. However, because of this extra load brought about by unintelligibility, we deemed it appropriate and informative to conduct a separate, additional analysis of CB instances, which followed a different coding protocol. The reasons for this decision are revisited later in this section, as well as in sections 3 and 5.1.

Our working definition of a communication breakdown, given above, is a broad one. It incorporates cases which other scholars may also term "misunderstanding" (e.g., Mauranen 2006) or "miscommunication" (e.g., Dascal 1999), and also includes non-understanding (as done e.g., in Jenkins 2000). In a similar vein, we adopt an all-embracing definition of intelligibility. We follow Bamgbose (1998: 11) in taking it to mean "a complex of factors comprising recognizing an

expression, knowing its meaning, and knowing what that meaning signifies in the sociocultural context". Other linguists, however, make a distinction between those three aspects, labelling them as intelligibility, "comprehensibility", and "interpretability", respectively (e.g., Smith & Nelson 1985; McKay 2002).

In general, breakdowns are more common in NS-NNS than in NS-NS dyads due to the fact that NSs and NNSs "may have radically different customs, modes of interacting, notions of appropriateness and, of course, linguistic systems", which renders them "multiply handicapped" in interactions with one another (Varonis & Gass 1985: 327, 340). NS discourse may present processing challenges to the NNS interlocutor, for example by virtue of showing insufficient accommodation to the needs of the latter. Embracing these needs ideally means avoiding "slang, opaque idioms, rapid speaking rates, and culture-specific references" (Trudgill 2005: 82). On the other hand, among the major difficulties inherent to NNS output, one can invoke their insufficient mastery of the linguistic system (however that mastery, or indeed the linguistic system to be mastered, is defined), which may result in what Varonis & Gass (1985: 334) term "noise" in the speaker's utterance, produced by, for instance, accent or ungrammaticality. This, in turn, will often act as a trigger of a corrective episode.

The main rationale behind examining CF instances, with special attention given to L2-speech-induced communication breakdowns, is their potential for carrying pedagogical implications. Establishing the types of non-targetlike linguistic structures that tend to invite corrective feedback, especially if they contribute to communication breakdowns outside of the classroom, could inform L2 teaching priorities (here: for English and French). Assuming that rendering L2 speech communicatively effective is a top priority in most types of L2 instruction, attempts to identify the types of errors which compromise intelligibility hardly need vindication. On the other hand, certain non-standard productions do not lead to communication breakdowns but still trigger corrective feedback from the interlocutor, as shown in example (1). They may therefore be argued to also merit special pedagogical attention, although possibly less so than those non-target forms that are communicatively more salient. In analysing both types of sequences, however, the key problem we have had to address is data ambiguity. This stems from the fact that the interlocutors' intentions and motivations – unlike the literal meaning of their utterances – are often far from evident, even when considered within a larger context and supplemented by visual cues. This is compounded by the conversational nature of our data, where a simple confirmation check or a genuine question on the part of the NS may easily be misinterpreted as an interrogative recast (i.e., CF) or even a sign of non-understanding.

The following sections will give more details on the SITAF corpus (Section 2) before offering a literature review (Section 3) and presenting our analyses of overall CF and then CBs found in the data (sections 4 and 5, respectively), homing in on various dilemmas we have encountered while coding and subsequently interpreting our findings (Section 6).

2 The SITAF tandem corpus

The SITAF corpus is a bilingual database of tandem exchanges collected at the University of Paris 3 – Sorbonne Nouvelle in 2013. The corpus, described at length in Horgues & Scheuer (2015), consists of around 25 hours of audio- and video-recorded, face-to-face interactions held by 21 pairs of native-French speaking and native-English speaking partners. The participants were all students at our university, aged between 17 and 22, none of whom were balanced English-French bilinguals.

To the best of our knowledge, no video corpus of spoken, face-to-face tandem exchanges had previously been compiled. The available language tandem corpora have mainly focused on written L2 production and/or made use of technology-mediated forms such as e-tandem or, more generally, telecollaboration (e.g., Ware & O'Dowd 2008; O'Dowd & Ritter 2006; O'Rourke 2005). Filling this gap, especially in terms of collecting real-time spoken material illustrating English/French tandem exchanges,¹ was the overall primary objective behind the SITAF project. From that perspective, we believed the SITAF corpus would have three principal assets. Firstly, with the added benefit of video recording, it allows for multimodal studies of real-time interactional phenomena, including non-vocal ones such as gestures or facial mimicry. Secondly, it provides a rich and comprehensive collection of speech data: apart from the NS-NNS exchanges, which constitute its crucial part, the corpus also contains L1-L1 data in both English and French, produced by the same set of participants. Each speaker therefore contributes three types of speech: L1 in an interaction with a fellow NS, L1 in an interaction with a NNS, and also L2 use in an interaction with a NS. In addition, the speech tasks and corresponding speech styles are varied, ranging from semi-spontaneous conversation (both narrative and argumentative) to text reading (see below). Thirdly, the SITAF database is longitudinal, allowing for the observation of a potential evolution in a learner's linguistic output and/or partners' interactional strate-

¹The choice of the English/French combination stemmed from the fact that the project was led by researchers from the English department at our university, i.e. a French university.

gies, possibly affording more insight about language and communication development during tandem learning.

The candidates for the SITAF project were all recruited on a voluntary basis, as part of an optional programme of autonomous tandem exchange run throughout the second semester of the academic year 2012/2013. The recruitment was performed with the help of an online questionnaire, which aimed to gauge – through self-assessment – dimensions such as their linguistic background (all languages spoken), level of proficiency in English (for NSs of French) or French (for NSs of English), as well as matters like interests and preferences regarding potential conversation topics and special requests as to their ideal tandem partner. Aside from the researcher's need to establish the participants' profiles with a view to interpreting and qualifying future findings, this information was deemed vital in the context of the pairing-up task, performed by the SITAF team members prior to the introductory meeting, during which all participants met their suggested partners. 45 tandem pairs were formed in this way. Of those, 25 subsequently took part in the first recording session, and 21 went on to attend the second session three months later (that is, completed the entire cycle). It is the data obtained from those 42 speakers that makes up the central core of the corpus and that the present study is limited to. The remaining pairs either did not respond to our invitations to the recording studio, or were unable to participate.

The 21 native French-speaking students (subsequently coded F01 to F21, to ensure anonymity) were English language majors for the most part, with a self-assessed level in L2 English of 7.2/10 for their mean proficiency and 6.8/10 for oral expression in particular. The 21 English-speaking students (coded A01 to A21) studied various disciplines and came from various Anglophone countries (United States, Canada, United Kingdom, Ireland). They self-assessed their level in L2 French as 6.9 out of a maximum of 10 for the mean proficiency and 6.6/10 for oral expression.² The above scores were crucial factors informing our pairing-up decisions, as we aimed to match candidates with similar self-assessed L2 proficiency levels, even if those were not necessarily expected to be very accurate reflections of the participants' actual abilities. The significance of proficiency pairing is acknowledged in the research on collaborative learning (e.g., Storch & Aldosari 2012), although tandem learning, again, does not present a standard case here, in that two different L2s are involved in each pair. Still, our rationale

²Out of the five self-evaluations (oral expression, oral comprehension, written expression, written comprehension, mean score), oral expression was the only score where the difference between the two speaker groups – native English vs. native French – reached statistical significance, with the latter group reporting a higher level of L2 proficiency ($p < .05$).

8 Potential pitfalls of interpreting data from E-F tandem conversations

was that a stark imbalance between the partners' L2 skills might lead to unnecessary frustrations on the part of one or both participants, which we preferred to avoid. All 21 native French speakers were female, whereas the Anglophone group consisted of 16 female and 5 male members.

The speakers were recorded on two occasions – in February (Session 1) and May 2013 (Session 2). Needless to say, in keeping with the principle of autonomy (e.g., Brammerts & Calvert 2003), the tandems were free to meet as often as they wished outside of the recording sessions. The questionnaires that the speakers filled out after completing the entire recording cycle suggest that the average tandem had met 12 times over the 3-month period in question, in line with the programme's recommendations for weekly conversations autonomously planned by the tandem partners. However, the individual numbers ranged from two to 23 meetings, thus pointing to substantial variation among the pairs.

Predictably, one of the dilemmas we faced while developing the experimental design was how to strike a balance between spontaneity and homogeneity of the data to be sampled. The latter quality is particularly valuable in the case of pronunciation studies – the main language area the authors specialise in – where having control over the phonemic makeup of the utterances greatly facilitates the researcher's subsequent analyses. As a result, we settled on three types of collaborative tasks, which came with a uniform set of written instructions in the participant's L1, to make sure each pair followed the same protocol. Two of them were communication activities: *Liar-Liar* (Game 1; expected to elicit a narrative style and the most spontaneous speech out of the three tasks) and *Like Minds* (Game 2; debating style with a pre-determined topic). The last one was a partially monitored reading task. In Game 1, the L2 learner had to tell a story containing three lies that the native-speaking partner had to identify by asking questions. In Game 2, both participants had to give their opinion on a potentially controversial subject – e.g., "Prisoners should not have the right to vote" – before assessing the degree of like-mindedness (in other words, convergence of opinions) between them. As regards potential metalinguistic interventions on the part of the NS during Game 1 and Game 2, the guidelines given to the Anglophones read: "When your partner speaks in English, let them do so as much as possible. However, feel free to help or correct them if they can't find the right word or expression, or if you think what they are saying needs correcting". The French participants were instructed accordingly about the French tasks. The text used for the reading task was "The North Wind and the Sun" ("La bise et le soleil" in French), which is a reference text in studies on phonological variation. The NNS speaker read the text twice. The first reading encouraged help and feedback from the NS part-

ner (hence *monitored reading*), whereas no interruption was supposed to occur during the second reading, which immediately followed.

We insisted on separating and balancing the use of the two languages, in that the entirety of the spontaneous tasks – i.e., both Games – had to be performed first in English and then in French, or vice versa.³ For the most part, the two recording sessions followed the same pattern, outlined above. However, Session 1 actually started with L1-L1 interactions (Games 1 and 2) before moving on to the L1-L2 exchanges, and Session 2 ended with text reading performed by the NS partner. Also, care was taken that in Game 2 each tandem discussed a different topic in each recording session and in each language condition. This was meant to ensure the novelty of the opinions being confronted, and therefore to promote a higher level of engagement of the participants.

Since the central focus of this chapter is data ambiguity, it is the findings from the communicative Games 1 and 2 that are discussed in the following sections. The reading task, being scripted and therefore essentially lacking in spontaneity, was deemed much less suitable for this type of analysis.⁴ It could be described – following Long (1991) – as a focus-on-forms task, since it is the linguistic (and more specifically, the phonetic) form of the learner’s output that the activity almost exclusively focuses on. Game 1 and Game 2, on the other hand, fall under the category of focus-on-form tasks, which are primarily concerned with communication, but which are nevertheless punctuated by the participants’ attending to linguistic issues (i.e., engaging in a Language-Related Episode (LRE)). This ties in with Loewen’s (2018: 2750) definition of focus-on-form practices as ones consisting of “primarily meaning-focused interaction in which there is brief, and sometimes spontaneous, attention to linguistic forms”. The following section presents a brief review of some of the relevant studies on LREs, and more specifically, CF. Consequently, it also outlines our framework for studying communication breakdowns.

3 LREs, CF and communication breakdowns

In this section, we attempt to clarify the relationship between Language-Related Episodes (LREs), corrective feedback (a subset of LREs), and communication breakdowns (to a large extent, a subset of CF).

³In principle, we alternated between English-first and French-first (i.e., every other tandem started in English and the others in French).

⁴Naturally, this is not to say that the interpretation of the reading data does not pose problems of its own (some of those are discussed in Horgues & Scheuer 2014). However, the nature of the task and the issues associated with it is sufficiently different to warrant a separate treatment.

Following Swain & Lapkin's (1998: 326) classic definition, an LRE is understood to be a part of interaction during which the participants "talk about the language they are producing, question their language use, or correct themselves or others". A few years earlier, the same authors offered a slightly differently worded definition, which explicitly stated that each LRE "is related to a *problem* the student had with the production of the target language" (Swain & Lapkin 1995: 379, italics added), thus pointing to the fact that the main driving force of such episodes is a potential gap between the target form and the form actually produced, or the absence of the latter. It is this original definition that is the default in the present chapter. So far, LREs have often been studied in classroom settings, during collaborative tasks performed by learners having the same L2 (e.g., Storch 1998; Basterrechea & García Mayo 2013; Basterrechea & Leiser 2019). As for studies of LREs in expert-novice interactions, which may not be the default LRE experimental context but which are of direct relevance in this chapter, CF either features prominently (e.g., Ballinger 2012) or is all but equated with LREs (e.g., Ware & O'Dowd 2008). Ballinger (2012: 79–80) clarifies that "all CF can also be categorized as LREs" but, in her study, giving and receiving CF (as well as partner-directed questions) were analysed separately from LREs in general "because they were deemed the most important for the promotion of collaborative interaction and for reciprocal language learning". This approach is replicated in our studies. Our decision to carry out a separate analysis of communication breakdowns, even though CBs in the context of L2 speech largely fall under the CF umbrella, follows the same logic. Communication breakdowns provide data deemed invaluable for the understanding of how NS-NNS (un)intelligibility works, and – consequently – for reciprocal communicative language learning and teaching practices. We therefore believe they occupy a pre-eminent position within CF episodes.

The importance of supplying CF, in one form or another, and the ways in which it can be beneficial to the L2 learner have been a recurrent theme in second language acquisition (SLA) literature, Nassaji & Kartchava (2017; in press) being examples of entire volumes devoted to the subject. Ellis (2017: 4), drawing on an early article by Hendrickson (1978), groups the key aspects of CF, both in terms of teacher experience and research findings, under the following five headings:

1. Should learners' errors be corrected?
2. When should learners' errors be corrected?
3. Which errors should be corrected?
4. How should errors be corrected?
5. Who should do the correcting?

The bulk of the findings from CF studies fall under one or more of the above categories, and they are briefly discussed in the following paragraphs.

First, the question as to whether learners' errors should be corrected at all. As Ellis (2017) points out, contrary to some suspicions expressed by the advocates of certain language teaching methods such as the Audiolingual Method or the Natural Approach,⁵ there has now been a wealth of research showing that CF does assist L2 acquisition (e.g., Li 2010; Lyster & Saito 2010). As for the theoretical grounding of the benefits of CF, Saito (in press) attributes them to CF's "ability to promote learners' awareness, noticing and understanding of linguistic form, especially when using their L2 for meaning conveyance".

The question of the timing of CF has been scarcely investigated, although immediate feedback appears to be preferred on theoretical grounds, for instance by virtue of providing the learner with a window of opportunity during which to map a specific form onto the meaning conveyed (Doughty 2001 cited in Ellis 2017: 7).

To date, most studies have investigated CF provided by teachers on L2 morphosyntax and vocabulary (Lyster & Ranta 1997; El Tatawy 2002; Mackey 2006; Lyster et al. 2013; Kartchava 2019). Meanwhile, some other studies have pointed to pronunciation and vocabulary CF being more noticeable for learners than morphosyntactic CF, which was found to be less likely to lead to uptake (Mackey et al. 2000; Saito & Lyster 2012; Saito in press). Learner uptake is understood, following Lyster & Ranta's (1997) definition, as "a student's utterance that immediately follows the teacher's feedback and that constitutes a reaction in some way" to that feedback.

As regards the question of how errors should be corrected, several typologies of CF strategies have been proposed with a view to establishing what the most frequent and the most effective type(s) are. Generally, researchers have classified these strategies on a continuum ranging from the most explicit to the most implicit CF (see Sheen 2006 for a relevant discussion). Lyster et al. (2013) distinguish strategies that offer negative evidence only: prompts (which include – with increasing explicitness – clarification requests, repetition of learner error, paralinguistic signals, elicitation and metalinguistic clues) from those which offer both negative and positive evidence: reformulations (including – with increasing explicitness – conversational recast, didactic recast, explicit correction, explicit correction + metalinguistic explanation). Lyster & Ranta's (1997) classic defini-

⁵For example, the Audiolingual Method believed in "strict control of learner output, thus removing the need for CF, which was viewed as a form of punishment that can inhibit learning" (Ellis 2017: 4).

tion of a recast describes it as a strategy involving “the teacher’s reformulation of all or part of a student’s utterance, minus the error”. In terms of the relative effectiveness of CF techniques (in the sense of being beneficial to L2 learning), studies have so far yielded variable results. Kartchava & Ammar (2014), for example, set out to determine the relative effectiveness of recasts, prompts and combinations of the two, in terms of both CF noticeability and L2 learning. The study was conducted on selected morphosyntactic structures in a classroom setting. Rather predictably, recasts proved to be the least noticeable of the three, although no significant differences across the CF types were found in terms of learning outcomes. Sato & Loewen (2018: 514) comment on previous studies by observing that “at least in the classroom setting, output-prompting corrective feedback has been found to better facilitate L2 development, compared to input-providing corrective feedback”, e.g. recasts. This was corroborated by their own study, where the former category was found to be more effective than the latter. There, however, the effectiveness of both types of CF was mediated by the linguistic structure concerned. The superiority of the output-prompting over the input-providing type was found to be statistically significant only in the case of the more perceptually salient structure under consideration (possessive determiners versus third-person singular *-s*). Saito (in press) summarises the available research by stating that explicit/output-prompting feedback may be particularly effective in a classroom context, whereas in laboratory settings, “where L2 learners can receive individualized attention from their interlocutors, all CF techniques seem to be equally salient and effective”.

The last question on Hendrickson’s (1978) list – “Who should do the correcting?” – has not as yet received a straightforward answer either. Even though most CF studies to date have looked into CF provided by teachers, the benefits of peer feedback – in the sense of learner-to-learner exchanges – have been receiving more and more attention in recent years (e.g., Adams 2007; Sato & Lyster 2012; Sato 2017). Sato (in press) observes that learners feel more comfortable working on a task with their peers compared to with the teacher or a native speaker. This is conducive to producing a higher amount of output, which in itself is of benefit to L2 learning. On the other hand, learners may not feel comfortable providing CF to their classmates, since that may be considered a socially inappropriate, face-threatening act (e.g., Foster 1998; Ballinger 2015). What is more, even if peer CF is provided, its quality may be problematic and its quantity insufficient. Mackey et al. (2003), for example, draw attention to its possible shortcomings on both counts (quality and quantity), even though other researchers have obtained results suggesting a longer lasting effect of peer, as opposed to teacher, CF (e.g.,

Sippel & Jackson 2015). Importantly, Sato (2017) links the nature and effectiveness of peer CF to the extent to which the social dynamics between the peers are collaborative⁶. If the learners fail to construct a collaborative relationship, there may be “social awkwardness in providing feedback, and embarrassment over being corrected by their peers” (Sato 2017: 27).

To conclude the above discussion, we may state that CF is a highly complex issue where no overall ideal strategy might necessarily be identified (e.g., Lyster et al. 2013), even though its benefits to L2 learning have been well documented, and “[l]earners almost invariably express a wish to be corrected” (Sheen & Ellis 2011: 606). This review of studies on corrective feedback serves as a backdrop against which to view research into the CF found in the SITAF database, summarised in Section 4.

4 Research on CF in the SITAF corpus

This part presents the main results of the CF analyses carried out on the SITAF data so far, before moving on to some of the key methodological challenges encountered in the process. More specifically, Section 4.1 expounds the criteria and parameters according to which we coded each corrective episode, and Section 4.2 offers the main findings in relation to those same parameters. Both sections set the scene for the discussion of coding issues that follows in 4.3.

4.1 Computing and coding CF episodes

Using the definition given in Section 1 as a starting point, we employ the term CF to refer to the verbally expressed negative evidence given by the NS participant to their NNS tandem partner during the recorded interactions. Naturally, “correcting others” (as per Swain & Lapkin’s 1998 definition) is just one among several possible types of LREs to be explored in the SITAF corpus, but it is the one that this chapter focuses on, with special attention given to communication breakdowns. Language-related episodes revolving solely around positive feedback (acceptance or acknowledgement of the correct form produced by the learner) or self-corrections, for example, are not discussed here.

The video recordings corresponding to the two communicative activities – Game 1 and Game 2 – in both recording sessions and in both languages were

⁶In his discussion of the social dynamics between peers, Sato draws on Storch’s model of dyadic interactions classified along the dimensions of equality and mutuality (e.g., Storch 2002).

8 Potential pitfalls of interpreting data from E-F tandem conversations

examined for the occurrences of CF. The simultaneous visual and auditory analysis was conducted by the two authors, who split the work but consulted one another (and, if necessary, other team members) about difficult or dubious cases and subsequently reached a consensus. Each CF occurrence thus identified was annotated and coded according to at least four parameters. This coding protocol expanded on some of Hendrickson's (1978) categories discussed in Section 3, notably aspects 3 and 4: "Which errors should be corrected?" and "How should errors be corrected?". The four parameters were the following:

1. CF focus: morphosyntax, pronunciation, vocabulary, or mix of any of the three
2. CF strategy (see below)
3. Presence or absence of CF request: whether the feedback was somehow solicited by the learner (explicitly or implicitly) or it was offered spontaneously
4. CF uptake: total, partial, failed or none.

Furthermore, CF episodes contributed by selected tandem pairs were also coded according to the multimodal resources employed by the participants, for instance types of gestures or specific vocal non-verbal content (hyperarticulation, rising tone, etc.).

As regards CF strategies, we have simplified and tailored Lyster et al.'s (2013) typology, presented in Section 3, to better fit the context of tandem exchanges. Importantly, we do not make use of categories such as elicitation or repetition of the learner's error, which are absent from our peer-to-peer interactions. These CF strategies seem to be restricted to teachers' corrective style and avoided by tandem participants possibly because they reinforce the asymmetry between the two partners.

Consequently, we have distinguished three basic CF strategies in our analyses:

1. Explicit comments, where the NS provides metalinguistic information explicitly, as in example (3) below, where the native English speaker is explaining to her unconvinced French partner why she made a mistake when she used *sunbath* as a verb:

(3) NNS: 'Cause you say bath.

NS: Right, bath, but then to bathe oneself, so sunbathe.

2. Clarification requests, as in example (4) below, where the NS reacts to her American partner's use of the word *flip-flops* in the middle of a French sentence:

(4) NS: Les flip-flops, c'est quoi? (French)
'What are flip-flops?' (*flip-flops* not being a French word)

3. Recasts, e.g.:

(5) NNS [talking about a past event]: And I miss my plane.
NS: You missed your plane?
NNS: Yeah, yeah.

Such input-providing corrections as recasts – unlike output-prompting strategies – effectively supply the novice “with reformulations in response to their errors, thereby providing positive evidence, that is, linguistic information about what is allowed in the target language” (Sato & Loewen 2019: 32). Yet, their *raison d'être* is the provision of *negative* evidence in the sense of signalling the incorrectness of the learner's output.

4.2 Summary of the main findings

The most comprehensive report to date on the corrective feedback found in the SITAF database is provided in Scheuer & Horgues (2020). The summary offered below is organised according to the parameters outlined in Section 4.1, i.e.: CF focus, strategy, presence or absence of request, uptake and multimodality.

All in all, we have identified 492 CF instances in the approximately 15 hours of conversational exchanges held as part of Game 1 and Game 2 in both recording sessions. Of those, 156 were found in the English, and the remaining 336 (i.e., over twice as many) in the French part of the data. The primary focus of the corrective interventions is vocabulary: about half of the CF instances found in the conversations – 52.5% in English and 49% in French – target missing or incorrectly used words, expressions or collocations. In the case of the French data, vocabulary errors also include the wrong grammatical gender, as in:

(6) NNS: Pour mes, ma Noël... (French)
For my.PL my.F.SG Christmas
NS: Ton Noël.
your.M.SG Christmas
NNS: Je veux parler de ma Noël.
I want to talk about my.F.SG Christmas

8 Potential pitfalls of interpreting data from E-F tandem conversations

The runner-up category in English was pronunciation, which accounted for 20% of all CF instances (e.g., the NS recasting the incorrect stress pattern in his partner's rendition of *pri'soners), followed by morphosyntax with 12.5%, as in:

- (7) NNS: Well, it depends of... the crime.
NS: On the, yeah.
NNS: On? It depends on the crime.

In French, morphosyntax ranked second, with 19% of hits, compared to 15% garnered by pronunciation. The remaining corrective episodes (15% in English and 17% in French) were classified as having mixed focus, as they revolved around learner utterances that were erroneous in more ways than one. These will be discussed further in Section 4.3.

As regards the corrective strategies employed by the SITAF participants, by far the most common one was recast, which accounted for 84% of CF in the English, and 89% in the French conversations. The remaining cases were split almost equally between explicit comments and clarification request.

Feedback was solicited by the learner roughly as often as it was not (i.e., the NS intervened unprompted in nearly 56% of the English cases, and just over 47% of the French ones). An example of solicited CF, which took the form of an explicit comment, is given in (8):

- (8) NNS: And after I celebrated the happy new year, you know [...]
the new year, not the happy new year, yeah?
NS: New year [head nod]. "Happy New Year" is what you say!

The two extremities of the uptake spectrum – total uptake and no uptake – jointly account for nearly 90% of all CF episodes in the two languages. However, there is a sharp difference between the English and the French conversations in terms of the relative share of each of the most frequent categories. In English, total uptake (shown in example 7) occurred in just 36.5% of cases, whereas no uptake followed 52.6% of the corrective interventions, with the remaining cases representing either partial or failed uptake. The French figures are almost identical but in reverse order: 52.4% for total, and 36.9% for no uptake. Example (6) illustrates the latter category: the recast of 'Noël', with the gender-appropriate possessive determiner, does not affect the NNS's subsequent utterance in any way.

Finally, corrective episodes proved to be highly multimodal activities. 94% of the CF occurrences studied were multimodal (i.e., combined verbal, vocal and visual resources). The remaining 6% were verbal and vocal only (Debras et al. 2015).

4.3 Some methodological issues encountered

The results presented in the previous section are the fruit of an analysis that was rich in methodological challenges. Whether or not a given episode constitutes an occurrence of CF is not always a straightforward matter. Furthermore, the exact strategy employed by the expert and the focus of their corrective intervention can also be hard to determine. Below we discuss and give examples of some of the challenges we faced while coding the data.

As could be expected from conversational data, the exact nature and purpose of the participants' output is not always easy to establish, and "the categorization of an utterance can be ambiguous since researchers are not privy to the speakers' intentions" (Ballinger 2015: 44). Example (5), where the NS produced an echo question albeit with the correct(ed) grammatical tense, serves as a good illustration of the most fundamental methodological issue we have been grappling with when coding CF instances in the SITAF database: deciding whether we are dealing with a corrective intervention in the first place. Recast is, by definition, discreet, indirect and non-threatening, not to mention easy to dispense. As such, it is particularly well suited to the tandem type of peer-to-peer interaction, where neither partner tends to particularly want to reinforce their short-lived dominant position. Therefore, it comes as no surprise that the vast majority of what we ultimately considered to be CF endeavours on the part of the NS (including example 5) were carried out by means of recast. However, precisely because of this discreetness, recasts may easily be misconstrued as the interlocutor's innocuous contribution to the activity. The problem is exacerbated by the conversational nature of our data (i.e., the fact that the participants engage in a genuine exchange of stories and ideas, where backchannelling, confirmation checks, repetitions and reformulations are common in both L1-L1 and L1-L2 interactions). The issue – although viewed from the perspective of the addressee of the hypothetical CF – has been highlighted by a number of researchers. Carpenter et al. (2006: 209), for example, observe that "recasts might be ambiguous to learners; that is, instead of perceiving recasts as containing CF, learners might see them simply as literal or semantic repetitions without any corrective element". This ties in with Sato & Loewen's (2019: 38) definition of a conversational recast as any teacher response to an error that includes "the correct linguistic form without any emphasis". In the context of tandem exchanges, where the NS partner is not a teacher but rather an empathetic peer, the remedial element of a conversational recast may be all the easier to miss.

In the short tandem exchange cited in (5), the NS's question, which we labelled as a recast, might have been nothing more than a confirmation check or a com-

miserative reaction triggered by the news of her partner having missed her plane. Those are particularly valid assumptions in the case of a communicative task like our narrative Game 1 (during which the exchange took place), where the NS listener was meant to make a mental note of the details of the NNS partner's story and was therefore expected to try to verify the information given in that story. The problem is also evident in the following example, where the French speaker erroneously pronounces the word 'castle' with a [t]:

- (9) NNS: My aunt organised a big party in a castle,
 NS: OK
 NNS: with all the family, with the cousins...
 NS: In, in a castle, you said?
 NNS: Yeah, a little casTle [gesture representing a castle] ...
 NS: OK [smiles].

Clearly, the NNS takes her partner's question at face value, and subsequently provides a gesture-enhanced confirmation – “yeah, a little casTle” – where she repeats her original error.

Needless to say, if recasts are ambiguous to the potential recipients, they may also be ambiguous to the researchers analysing the exchanges, because the researchers are unable to tap into the speakers' mindset. What they do have access to, however, is the remainder of the conversation, which may, retrospectively, shed more light on the NS's intentions. For example, knowing that the American speaker quoted in (9) tried to elicit the word *castle* from his partner later in the conversation gives us extra reassurance that his “in a castle, you said?”, uttered 3 minutes earlier, was indeed an interrogative recast rather than a genuine question. Still, our decisions to classify many cases like the above as CF attempts, although often reinforced by the study of the surrounding context, are nevertheless not entirely unquestionable. Ultimately, whether the NSs in (5) and (9) were actually trying to correct their partners' pronunciation and/or syntax, rather than simply making sure they had correctly understood the discourse, will never be established beyond doubt. After all, it is not every day that one misses a plane or one parties in a castle.

On the other hand, such potential mistakes, where a CF label was applied to an utterance devoid of corrective intention, may have been offset by a number of instances where the opposite error may inadvertently have been committed. The error would have consisted in actual recasts being miscoded as conversational turn-taking. In particular, this was likely in cases where the NS repeated all, or some, of the NNS's utterance word-for-word, phoneme-for-phoneme, as in the following exchange:

- (10) NNS: Il y avait trois. (French)
‘There were three’.
NS: Trois, d’accord.
‘Three, all right’.⁷

Much as the above looks like an innocuous repetition, confirmation or simply an effort to maintain the conversation flow, one cannot rule out the possibility that the expert was, in fact, recasting a suprasegmental or sub-phonemic detail of the non-native pronunciation (e.g., the /ʀ/ in ‘trois’ /trwa/, which the NNS all but deleted) that they judged worthy of discreet correction. Technically, the NSs’ interventions like the one shown in (10) do satisfy our broad definition of recast – repeating all or part of the novice’s utterance minus the (pronunciation) error – even though the presence of actual corrective intention is far from evident.

Even if a CF instance has been rightly acknowledged as such, this does not automatically mean its classification – both in terms of the exact CF strategy employed and of its intended focus – is a clear-cut matter. Both aspects may be complex to interpret since a correction may involve multiple moves (e.g., a recast followed by a clarification request) and may target various linguistic levels (phonology, lexis, morphosyntax) at once. While coding errors relating to a CF strategy can be argued to be relatively inconsequential, failure to identify the linguistic trigger of the NS’s remedial reaction could potentially skew certain pedagogical implications that the researcher may wish to glean from the CF analyses. The validity of such implications hinges on determining a causal link between the specific characteristics of the NNS’s output and the NS’s corrective intervention. Assuming that the experts are relatively likely to intervene when the novice’s mistake is prominent in terms of compromising communication (see Section 5) or violating a norm that they consider important, being able to dissect the nature of that mistake is of great potential value to language teachers. Namely, it helps pinpoint the types of interlanguage errors that bother NSs more than others, which might, in turn, inform teaching priorities. Given that not every aspect of the L2 system can possibly be accorded the same amount of time and attention in the L2 classroom, being selective about what to teach first and foremost is a sheer necessity. Even the most committed language instructor has therefore to make choices, which may well be guided by the perceived gravity of non-target forms. If so, examining some of the factors which might potentially determine this gravity seems like a worthwhile endeavour.

⁷In fact, the correct version of the NNS’s utterance would have been ‘il y *en* avait trois’ [‘there were three of them’], but this syntactic inaccuracy clearly did not bother the French NS enough to rectify it.

This endeavour is certainly not without its problems. One of the pitfalls inherent in attempting to establish relative error gravity is confusing different dimensions, such as the dimension of accuracy with that of communicative effectiveness. Namely, it may be tempting to conclude that an error which may potentially lead to a communication breakdown is somehow more ‘erroneous’ than an error that does not jeopardise intelligibility. Meanwhile, as pointed out by Pallotti (2009: 592), both types of error have the same impact on the accuracy of an utterance, since “a 100-word production with 10 errors not compromising communication is not more ‘accurate’ than a text of the same length with 10 errors hindering comprehension, but just more ‘understandable’ or ‘communicatively effective’”. Nonetheless, it can be argued that the types of errors that, for whatever reason, tend to command the NS’s corrective attention also deserve special pedagogical attention, at least in contexts where the learner is consciously oriented towards the NS model, as is indeed the case with tandem exchanges. To borrow Pallotti’s logic: a 100-word production with 10 errors not bothering the listener is no more ‘accurate’ than a text of the same length with 10 errors bothering the listener, but it nevertheless ranks higher than the latter on the dimension of acceptability, which is no small matter.

Since the vast majority of CF instances were performed by means of recast and the NNSs’ utterances were often incorrect in more ways than one, it comes as no surprise that the specific motive behind the NSs’ corrective intervention was not always evident, neither to the recipient, nor subsequently to the researcher. Example (11) is a case in point:

- (11) NNS: There were interior in leather [* [’li:ðər]].
 NS: Oh, the leather, oh, so there’s leather interior.
 NNS: Yeah, because in Ferrari there’s leather [* [’li:ðər]] in the car.
 NS: Yeah, yeah, it’s just what you said [gesture representing switching] you, just switch the words so it’s leather interior, not interior leather.

The American speaker first uses a recast (“leather interior”), then explicitly insists on the correct word order (“just switch the words”), then praises his French partner for eventually getting the order right, and subsequently resumes the flow of the conversation. However, throughout this episode, the French speaker persists in her erroneous rendition of the vowel, thereby only correcting her original output to “leather * [’li:ðər] interior”. It is impossible to know to what extent the pronunciation problem troubled the NS and whether it would have triggered his corrective intervention at all, had it not been accompanied by the syntactic error.

The fact that he chose not to revisit the NNS's utterance once the syntax had been fixed does not necessarily mean the wrong vowel was not considered an issue. Rather, the expert might have chosen not to overwhelm his partner with too many corrections directed at one short utterance, which might have been unhelpful to her L2 acquisition process (cf. Ellis et al. 2008, in the context of written CF).

Unlike the example above, there were numerous cases where a recast was not accompanied by an explicit comment. These instances were more enigmatic and therefore more problematic. This is illustrated by the following exchange:

- (12) NNS: We putted our ski.
NS: [nodding] Put your skis on.
NNS: And we ...

Here, the exact reason for the NS's remedial reaction – was it the wrong past tense verb form, the missing particle, or the missing plural marker? – is unclear, even though it stands to reason that the cumulative effect of the various issues may have been what incited the expert to intervene by recasting all the issues in one move. We believe that our decision to label such occurrences as having a 'mixed' focus (in the example above: a mix of lexis and morphosyntax) has the advantage of making our observations more objective, through minimising the need for the researcher's personal judgement and interpretation of what the speaker truly meant to correct while offering the correction.

Having presented the main results of our studies of general CF in the SITAF corpus, as well as the major problems encountered in the process of obtaining them, we will now turn to the other – related – research focus of direct relevance to the chapter: communication breakdowns.

5 Study of communication breakdowns

5.1 Communication breakdowns versus CF in general

The detailed study of communication breakdowns is a major thread of research to emerge from analysing the SITAF data on the back of CF analyses (e.g., Horgues & Scheuer 2018). Our working definition of a CB, provided in Section 1, encompasses all cases where the listener demonstrably has difficulty or is incapable of understanding the meaning of an utterance as intended by the speaker. This tallies with Mauranen's (2006: 128) definition of the term "misunderstanding", taken to denote "a potential breakdown point in conversation, or at least a kind

of communicative turbulence”. The reasons why CBs receive a separate treatment in our analyses are mostly to do with their particular communicative, and therefore potential pedagogical, relevance. The matter was already addressed in Sections 1 and 3, and is further clarified below.

The relationship between communication breakdowns and CF is not entirely straightforward. CF is often provided even though comprehension is not at stake, as examples (5–12) demonstrate. The reverse is also true: a communication breakdown in NS-NNS conversations may contain no CF overlay at all. This typically occurs when it is the NS’s discourse that is not understood, as in (13):

- (13) NS: On va pas le défendre; on va plutôt le sermonner. (French)
 ‘One will not defend him; one will rather lecture him.’
 NNS: Sermonner, qu’est-ce que ça veut dire?
 “Sermonner” [to lecture], what does it mean?
 NS: [explains the meaning of ‘sermonner’]

Clearly, there is no corrective intention behind the NNS’s clarification request, and the NS’s subsequent explanation only serves to provide the NNS with positive – rather than negative – evidence, even if the whole sequence uncovered a lexical gap on the part of the NNS. Another type of episode that can be potentially classified as an instance of communication breakdown but not of CF arises when the confused recipient is sending visual signals only. Non-verbal strategies, especially face expressions (frowns, squints) or shifts in gaze, may well be indicative of non-understanding. However, in accordance with our definition given in Sections 1 and 4.1, they do not, in or by themselves, count as CF in the present analysis. Finally, a CB may go undetected by either participant, even though it may be evident to an external observer. In such cases the NS expert, unaware of the true meaning of the NNS’s utterance, will not be able to provide correction. Only one such instance has been identified in the SITAF corpus. In one other episode, a CB very nearly went undetected: the French participant mispronounced the word “tuition” in “tuition fees” so that her NS partner misinterpreted it as “teaching fees” (i.e., teacher salary). A prolonged misunderstanding sequence follows (2’30 minutes long), where the two interactants run parallel but disconnected argumentations (what the students should pay to study versus how much the teachers should get paid). The problem finally gets resolved almost by accident, when the NNS makes a remark – “for studies” – which alerts the NS to the fact that she has been misunderstanding her partner all along:

- (14) NNS: Here, not a lot of people can afford 400 [euros] per year for, for studies, so [...]

NS: Oh, hang on [looks at the slip of paper with “tuition fees” printed on it]. Tuition fees, OK. [...] I thought, teaching fees. Instead of tuition fees. [...] I thought the teacher only gets paid 400.

NNS: [laughter] Ah, no, no [...].

Despite such divergences, there is a considerable area of overlap between communication breakdowns and corrective feedback, in that verbally signalled CBs arising in the context of NNS speech are largely a subset of CF instances. The ‘flip-flops’ example (4) illustrates this point: the very fact that the attentive NS (who, within a tandem setting, will by default be cooperative⁸) does not understand a lexical item will, in most cases, constitute negative evidence already: either the word itself is incorrect, or there is something wrong with the NNS’s rendition of it.

5.2 Computing and coding communication breakdowns

Needless to say, each CB occurrence which also constituted corrective feedback had previously been annotated according to the CF coding protocol (Section 4.1). In addition, each instance of communication breakdown identified in the data – whether or not it coincided with CF – was annotated and coded according to parameters such as:

- Whose speech got misunderstood: NS or NNS?
- Miscommunication trigger: morphosyntax, pronunciation, vocabulary or mix of any of the three?
- Who detected the misunderstanding: the main speaker, the interlocutor or both?
- Timing of detection: instantaneous (next turn) or delayed?

Comparisons were also made between Session 1 and Session 2.

Section 5.3 presents our main findings, along the lines outlined by the above parameters.

⁸The expectation that tandem partners should be cooperative and therefore willing to understand each other stems from one of the fundamental principles of tandem learning, i.e. reciprocity. This means “the reciprocal dependence and mutual support of the partners” (Brammerts 1996: 11). The listener’s willingness to understand the interlocutor, which is not always to be taken for granted, is a crucial factor in mutual intelligibility, highlighted by Chambers & Trudgill (1998: 4)

5.3 Summary of the main findings

Since one of our goals was to determine whose output was misunderstood in each case, we deemed Game 1 (storytelling) less suitable for this type of quantitative analysis, given that disproportionately more speaking time was naturally given to one of the participants (the storyteller). Quantifying communication breakdowns encountered in the course of that activity might therefore have skewed the overall results.

Having quantitatively analysed the data from the debating Game 2 (approximately 5h), we identified a total of 72 cases of detectable communication breakdowns in the two language conditions. Of those, 41 were found in the English and 31 cases in the French conversations. A total of 40 (55.6%) arose in connection with the NNS discourse, which means that in the remaining 32 (44.4%) cases it was the NS who was misunderstood or not understood. Vocabulary proved to be the main stumbling block when it came to processing NS discourse (21 cases out of 32; 65.6%), as opposed to pronunciation in the case of NNS speech (14 cases out of 40; 35%). In about two-thirds of occurrences it was the interlocutor (recipient) who signalled the communication breakdown, whereas in the remaining cases the problem was detected by both participants, roughly simultaneously. In keeping with the collaborative spirit of tandem learning, CB detection was largely instantaneous, occurring in the next turn (60 out of the 72 instances), occasionally delayed (11 instances), and missing altogether in just one case. The number of communication breakdowns dropped between the two recording sessions, from 39 to 33, although the difference was only statistically significant in the English conversations (from 26 to 15).

5.4 Some methodological issues encountered

It comes as no surprise that in many respects the methodological challenges posed by the identification and subsequent interpretation of communication breakdowns resemble those encountered while exploring general corrective feedback. Not only can it be difficult to pinpoint the exact cause of a CB, but it is also frequently impossible to determine with a fair degree of certainty that mutual comprehension was hindered.

Our analysis of CB instances has necessarily been confined to cases where a comprehension issue is somehow signalled. The problem is that such signals will inevitably vary in clarity and will therefore be more or less legible to the observer. Example (13) represented the clear end of the spectrum, as does (15), this time with the NS in the role of the non-understander:

- (15) NNS: Cela rend les gens plus seuls [*[sul]]. (French)
‘This makes people more lonely.’
NS: [at first, silence and blank face] Plus quoi?
‘More what?’
NNS: Seuls [*[sul]].
‘Lonely.’
NS: Ah, plus seuls [[sø1]]!
‘Oh, more lonely!’

The unambiguous clarification request on the part of the French participant (“more what?”) is a clear sign of her struggling to make sense of her NSS partner’s utterance.

On the other hand, the interlocutor may react to an utterance with signals so subtle as to leave the researcher in doubt as to their true significance, as in (16):

- (16) NNS: If it’s just uh..
NS: [keeps nodding]
NNS: if you’re just as thief [*[’vif]]
NS: [stops nodding]
NNS: who, who go to prison,
NS: [now nodding only slightly]
NNS: maybe you could? [...]
NS: Uhm. I don’t know where I stand on this.

Here it is the disappearance of non-verbal response on the part of the interlocutor that is suggestive of her confusion: she had been nodding for some time but stopped doing so upon hearing the [’vif] utterance. Yet again, the surrounding context is helpful: the NS’s subsequent verbal contribution, which is rather non-committal and does not build on her partner’s discourse, provides further support for this interpretation.

Apart from being economical with cues, the recipient may also be sending conflicting signals as to whether the tandem partner’s discourse has been understood. The following exchange can serve as an example of this coding dilemma:

- (17) French speaker:⁹ In French we say familiarly *une boîte à fric*, je sais pas si [...]
‘a scam, I don’t know if [...]
American speaker: OK [gazes sideways].

The conversation proceeds in English but, due to a lexical gap, at some point the French speaker resorts to an expression in her L1 (*une boîte à fric*). On a

purely literal level, her American partner seems to have no difficulty processing her output (he utters “OK”), but the non-verbal cues he provides tell a slightly different story: his tone of voice is hesitant and he gazes sideways, suggesting that, at least at that particular moment, the exact meaning of the colloquialism *boîte à fric* is unclear or the sudden language switch caught him off-guard.

In addition to the dilemmas outlined above – determining whether a communication breakdown did indeed occur, in the absence of tangible or consistent cues – the researcher is faced with the other major coding issue discussed in the context of CF: what brought the problem about. Identifying the linguistic triggers of communication breakdowns in NS-NNS interactions is potentially of even greater pedagogical importance than is the case with the remaining body of CF. After all, the primary function of language is communication. If that is jeopardised, one is justified in trying to eliminate the source of the problem before moving on to somewhat higher-level considerations potentially triggering CF provision, such as sounding aesthetically pleasing to the listener. Fortunately, the majority of communication breakdowns identified in the SITAF corpus leave the researcher in little doubt as to the linguistic source of the problem. Explicit information is often provided by the participants themselves during the relevant episodes. This is shown in example (18):

- (18) NNS: Et la cousine [*[kyzin]] de de mon père d'accueil n'a pas mangé du veau parce que elle a dit que... (French)
 ‘And the cousin of of my host father did not eat the veal because she said that...’
 NS: [confused facial expression at first] Ah, la cousine! [...] D'accord, j'avais compris la cuisine; la cousine, OK!
 ‘Oh, the cousin! [...] All right, I'd understood *the kitchen*, the cousin, OK!’

The NS makes it fairly clear that the problem was her partner's erroneous fronting of the first vowel in the word *cousine* (/kuzin/), which made her perceptually confuse it with *cuisine* (/kɥizin/). Unlike in (15), however, the NS offers her partner (and the researchers) the added bonus of an explicit comment on how exactly she misunderstood the NNS utterance. It is also worth noting that (18) is a representative example of how a communication breakdown may serve as a starting point of a corrective episode, a phenomenon alluded to earlier on in the chapter.

Despite the prevalence of relatively straightforward (in terms of the linguistic trigger) cases like (18), the cause of the communication breakdown was not always easy to pin down. As a result, in around 18% of instances (13 out of 72) we

ended up labelling the CB as being due to a combined trigger, for reasons similar to those mentioned in the context of the mixed focus CF occurrences, and with the same corollary of making our observations less informative than we might perhaps have wished them to be. The NNS in (19), using the expression *for per'petuity* to talk about prisoners sentenced to life in jail is one such instance:

(19) NNS:

And if you are in jail for... I don't know how to say that... for **per'petuity*

NS: What?

Not only is the phrase a calque from French (the word *perpetuity* not being used in this legal sense in English), but the NNS also mispronounces the word by stressing the second syllable instead of the third. The NS does not know what her partner is trying to say (“What?”), but a successful clarification attempt follows. This involves the French speaker first switching to her L1 (“perpetuité”) and then reformulating her initial proposition as “you’re gonna die in prison”. It is perhaps tempting to propose that the pronunciation issue played the key role in generating this temporary communication breakdown. After all, the English word *perpetuity* is not far removed semantically from the French term, so the NS would likely have gotten the idea had she been able to simply recognise the word, but this is impossible to verify.

Other instances where a CB is clearly evident present an even more complex picture in terms of their underlying cause(s). A case in point is a French conversation where what appears to be the keyword in the NNS’s discourse – *le lieu* [/lə ljø/], ‘the place’ (where she got sick) – is erroneously rendered as *[la ly]:

(20) NNS: *La lieu [*[ly]] où je suis tombée malade. (French)
the.FEM place where I got sick.

NS: [confused facial expression]

NNS: *La lieu [*[ly]].

NS: [la ly]???

The cumulative effect of the two issues, the wrong vowel and the wrong grammatical gender, ensures that the NS is at a loss as to what her partner means. She clearly does not understand, which is revealed not only by her confused facial expression and interrogative echoing of the offending sequence, but also by her unsuccessful attempt at paraphrasing it, which follows: “Ah! T’es pas tombée malade!” [‘Ah! You didn’t get sick!’. Matters are not helped by the American speaker’s subsequent use of a false friend – *la location* [‘the hire’] – in a bid to clarify the meaning of her original sentence, and the CB remains unresolved. Again,

it would be valuable to know whether one of the two mistakes involved in *[la ly] was actually more salient than the other, in the context of establishing error hierarchies. For example, it stands to reason that the mispronounced vowel, but not the incorrect gender, might have single-handedly pushed the native speaker over the line of non-understanding. If so, this would highlight the importance of attending to pronunciation details (here: the exact quality of rounded vowels) in the L2 French classroom, where correct gender assignment might receive considerably and undeservedly more pedagogic attention in comparison. What is worth noting is that no input-providing CF can be given in example (20), since the novice's utterance remains cryptic to the NS, while being clearly incorrect. The latter adopts a 'let-it-pass' strategy and chooses to simply move on, uttering a rather unconvinced "d'accord" ['all right'] in the process. Her American partner emerges from this episode none the wiser as to the grammatical gender and the phonemic shape of the word *lieu*, and she has reason to believe that her discourse, if not entirely accurate, was at least communicatively effective.

As shown in this section, dissecting the nature and identifying the exact trigger of a communication breakdown may be a highly challenging task, in ways that are similar to those previously discussed in the context of general CF instances.

6 Discussion and conclusion

This final section offers a summary and a further discussion of the methodological issues highlighted in the chapter, the ways in which we have tried to address them, as well as a conclusion hinting at potential future perspectives.

6.1 Methodological challenges encountered

As demonstrated in the chapter, coding LREs such as CF episodes and communication breakdowns occurring in semi-spontaneous NS-NNS interactions is no straightforward task. The participants' output and reactions can be both complex and ambiguous, often making it impossible for the researchers, or indeed the interlocutors themselves, to perceive and decode the speakers' intentions with a fair degree of certainty. This ambiguity has manifested itself particularly at the levels of identifying the interactional function of speech turns (e.g., teasing apart corrective sequences from conversational moves such as confirmation checks or topic continuations) and pinning down the triggers that led to a given CF or CB instance.

The added layer of complexity stems from the fact that the various functions, as well as the various triggers, may appear in combination with one another. Our

method of coding the two interactional phenomena – CF and communication breakdowns – acknowledges this complexity. This is reflected in our extensive use of the labels mixed and combined when more than one factor seemed to be at play. However, this kind of cautious coding will inevitably influence and, to some extent, constrain the interpretation of our findings. Since one of our stated objectives in studying CF and communication breakdowns has been to obtain data that could inform L2 teaching priorities, an optimal end result would be to provide unambiguous answers as to what types of non-native productions are likely to cause communicative turbulence. Intelligible speech is certainly one of the most highly desirable learning outcomes in any L2 classroom, which means that those non-target productions that tend to jeopardise intelligibility should perhaps receive the teacher’s attention before anything else. On the other hand, non-native output which simply triggers corrective feedback without hindering communication will probably rank considerably lower in that hierarchy, while still being more worthy of remedial action in the classroom than other types of inaccurate productions. For those reasons, mixed and combined CF/CB instances will be of more limited pedagogical relevance, since they will be harder to interpret in terms of specific remedial actions. Sequence (20) may serve as an example here: the fact that the non-target vowel in *lieu* is intermingled with the wrong grammatical gender, to some extent downplays the importance of either issue, as it is uncertain whether either of them would have triggered the communication breakdown on its own. Instead of possibly serving as a prime example of how incorrect vowel quality may single-handedly hamper intelligibility, this instance will lose some of its significance by feeding into the rather fuzzy mixed category.

Another problem with interpreting quantitative findings obtained from a group like ours is the fact that they are generalised across pairs that are far from homogeneous. The social dynamics between the two partners will naturally be slightly different within each tandem, and that will inevitably affect the way CF is dispensed and received and the way communication breakdowns are signalled and resolved. This means that the data contributed by different tandems may not always be directly comparable. As observed by Horgues & Tardieu (2015), certain SITAF participants are hyper-correctors and others are hypo-correctors, and there is no straightforward correlation between the level of L2 competence and the amount of CF received. For example, of the 336 CF instances we identified in the French section of the corpus, one participant (F11) contributed 52 (15.5%) cases of CF. On the other hand, two other French speakers produced just one instance (0.3%) each. Foster (1993: 25) highlights a similar issue in the context of her own study of collaborative tasks performed in an L2 classroom: “The range

in the individual scores is so wide, and the lack of participation by some students is so striking as to make statistics based on group totals very misleading”. This is another reason why the interpretation of such group observations in the context of gleaning pedagogical insights should be carried out with utmost caution. Due to idiosyncratic linguistic preferences and individual corrective styles, certain relatively minor inaccuracies might get overrepresented and therefore ascribed disproportionately more importance than they deserve, if they happened to fall on the over-sensitive ears of an over-corrector. F11 with her 52 corrective interventions is a case in point. On one occasion, she corrects a collocation (*heureuse* to refer to *période*) that her American partner has directly copied from the topic the pair was given in writing at the beginning of the task. The topic, which read “L’adolescence est la période la plus *heureuse* de la vie” (‘Adolescence is the *happiest* period of your life’), had previously been prepared and approved by her fellow native French speakers. In other words, the NNS gets corrected on something that for all intents and purposes is correct in L1 French, which might make this corrective intervention appear of little didactic value. On the other hand, it could be argued that it is precisely this sort of rather unexpected and unconventional results that make our findings most interesting. If one takes the participants’ perspective, one gets a chance to see what individual speakers treat as an error, or at least what sort of forms they find annoying and worth eradicating, on top of the “real” errors that one could identify and code by simply referring to handbooks and dictionaries.

6.2 Solutions adopted

In view of the fact that there seem to be no available studies of corpora of video recorded, face-to-face tandem interactions, we have had to grapple with challenges that have not necessarily been adequately addressed in the SLA literature. The frameworks previously developed for analysing LREs in L2 classroom settings do not entirely fit our context of expert-novice, yet peer-to-peer, interactions. Therefore, one of the basic steps we needed to take was to adjust the descriptive categories previously employed in CF studies, to better capture the specificities of our data. Crucially, ours was a setting where there was no need to account for corrective moves characteristic of teacher discourse, but where the roles of (relative) expert and novice within each conversation section were clearly defined.

While coding conversational data, which is invariably complex and ambiguous, the risk of exercising excessively subjective judgement is ever present. We

have endeavoured to minimise the role of this subjectivity through taking various measures. One basic and commonly adopted step, in addition to developing a detailed coding protocol, was to have the particularly challenging cases analysed by two or more team members. The multimodal nature of our data also provided further opportunity to objectivise our analysis. Supplementing the subtle – or even non-existent – verbal cues potentially signalling a communication breakdown with the vocal and visual cues (rising intonation, changes in speech rate, hesitation, facial expressions, gestures) proved extremely helpful in deciding on the most plausible interpretation of the sequences in question. Moreover, our aim was always to consider the CF/CB episodes within their larger contexts and therefore to benefit from the wisdom of hindsight. That meant looking not only at the turns immediately preceding and following the episode under scrutiny, but also taking into account the rest of the conversation. As our comments on examples (9) and (16) demonstrate, the participants' subsequent utterances may provide precious insights into their intentions and thus lend support – or not – to our hypotheses concerning the nature of the actions performed several seconds or minutes earlier. Needless to say, being able to watch the exchanges numerous times offers the researchers various opportunities for refining their hypotheses – another considerable advantage over the real-time processing that the participants themselves needed to execute. Making use of the rather vague labels mixed and combined when coding complex CF or CB instances represents further efforts on our part to minimise the effect of subjective judgement as to what the underlying linguistic triggers were. Such categories tally with the reality of L2 speech production and perception, where the intermingling of issues from various linguistic levels (phonetics, syntax, semantics) is the norm rather than the exception.

Lastly, there is an issue which is more or less implicit in the account of our CB data coding and which represents a challenge and a solution at the same time: the fact that we have only taken into consideration those communication breakdowns that are somehow overtly marked. As a result, a potentially large amount of covert communicative turbulence may have been left unaccounted for. Signalling non-understanding – just like giving CF – may be regarded as a face-threatening act. This means that certain participants may have refrained from sending distress signals as a deliberate strategy to prioritise fluid and friendly communication, in the hope that the meaning intended by their partners would get clarified later in the conversation. In a bid to keep our coding process as objective as possible, we chose not to speculate about – and, consequently, not

to quantify – such likely avoidance phenomena.¹⁰ This approach, however, undoubtedly affects the interpretation of our findings, in that our quantitative data almost certainly suggest that the participants misunderstood each other less often than they actually did. Yet again, though, it could be argued that this apparent shortcoming puts our results more in line with real life speech processing than might otherwise be the case, as – according to Keysar (2007) – speakers routinely believe that what they say is accurately understood by the addressee more often than it really is.

6.3 Conclusion

The SITAF tandem corpus captures conversational exchanges between various types of speakers in all their inherent complexity, multimodality and ambiguity. The fact that, by definition, tandem partners do not share an L1 makes matters even more complex and our data even more challenging to interpret, especially at the level of negotiation of form, than would presumably be the case with NS-NS dyads. Throughout the paper we have shown how we attempted to deal with the various aspects of data ambiguity, and how our decisions impact our conclusions.

The analysis of our corpus data could certainly be refined in the future, mainly by going further beyond the verbal and literal information contained in the participants' output. Research paths that could be explored in a bid to enrich our findings include issues like stance taking, power dynamics within individual tandem pairs, a variety of face-saving strategies employed, notions of politeness and appropriateness, affective and empathetic reactions, as well as task effects. In the event of compiling a new, similar corpus of tandem interactions in the future, data ambiguity could to some extent be reduced through employing a stimulated recall protocol (as done by Mackey et al. 2000, for example). This would enable the researchers to watch the recorded interactions with both participants, with a view to discussing their perceptions of the LREs they have just engaged in. Despite such measures, an element of ambiguity is still bound to remain when it comes to perceiving and interpreting real people's actions, intentions and emotions. There will therefore always be more to a database of human interactions than will meet the researcher's cautious eye.

¹⁰On numerous occasions it was tempting to engage in such speculations (for instance, when the participant was speaking very fast or indistinctly or their L2 production was extremely dysfluent).

Acknowledgements

The SITAF project was financed thanks to a research grant obtained from the Conseil Scientifique de l'Université Sorbonne Nouvelle (Projet Jeunes Chercheurs, 2012–2014). Part of the orthographic transcription and corpus finalisation was financially sponsored by Labex EFL program (ANR-10-LABX-0083) and Ircom/Ortolang. For their participation and support, we are very grateful to all the SITAF tandem participants, the SITAF team members, our research team (Sesylia, Prismes EA 4398) and the university's engineers.

References

- Adams, Rebecca. 2007. Do second language learners benefit from interacting with each other? In Alison Mackey (ed.), *Conversational interaction in second language acquisition: A collection of empirical studies*, 29–51. Oxford: Oxford University Press.
- Ballinger, Susan. 2012. *Towards a cross-linguistic pedagogy: Bilingual and reciprocal learning strategies in French immersion*. Montreal. (Doctoral dissertation).
- Ballinger, Susan. 2015. Linking content, linking students: A cross-linguistic pedagogical intervention. In Jasone Cenoz & Durk Gorter (eds.), *Multilingual education: Between language learning and translanguaging*, 35–60. Cambridge: Cambridge University Press.
- Bamgbose, Ayo. 1998. Torn between the norms: Innovations in world Englishes. *World Englishes* 17(1). 1–14.
- Basterrechea, María & María del Pilar García Mayo. 2013. Language-related episodes during collaborative tasks: A comparison of CLIL and EFL learners. In Kim McDonough & Alison Mackey (eds.), *Second language interaction in diverse educational contexts*, 25–43. Amsterdam: John Benjamins.
- Basterrechea, María & Michael J. Leiser. 2019. Language-related episodes and learner proficiency during collaborative dialogue in CLIL. *Language Awareness* 28(2). 97–113.
- Brammerts, Helmut. 1996. Tandem language learning via the internet and the international email tandem network. In David Little & Helmut Brammerts (eds.), *A guide to language learning in tandem via the Internet (CLCS Occasional Paper no. 46)*, 9–22. Dublin: Trinity College.
- Brammerts, Helmut & Mike Calvert. 2003. Learning by communicating in tandem. In Tim Lewis & Lesley Walker (eds.), *Autonomous language learning in tandem*, 45–59. Sheffield: Academy Electronic Press.

- Carpenter, Helen, K. Seon Jeon, David MacGregor & Alison Mackey. 2006. Learners' interpretations of recasts. *Studies in Second Language Acquisition* 28(2). 209–236.
- Chambers, Jack K. & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press.
- Dascal, Marcelo. 1999. Introduction: Some questions about misunderstanding. *Journal of Pragmatics* 31. 753–762.
- Debras, Camille, Céline Horgues & Sylwia Scheuer. 2015. The multimodality of corrective feedback in tandem interactions. *Procedia – Social and Behavioural Sciences* 212. 16–22.
- Doughty, Catherine. 2001. Cognitive underpinnings of focus on form. In Peter Robinson (ed.), *Cognition and second language instruction*, 206–257. Cambridge: Cambridge University Press.
- El Tatawy, Mounira. 2002. Corrective feedback in second language acquisition. *Working Papers in TESOL & Applied Linguistics* 2. 1–19.
- Ellis, Rod. 2017. Oral corrective feedback in L2 classrooms: What we know so far. In Hossein Nassaji & Eva Kartchava (eds.), *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*, 3–18. New York: Routledge.
- Ellis, Rod, Younghee Sheen, Mihoko Murakami & Hide Takashima. 2008. The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System* 36(3). 353–371.
- Foster, Pauline. 1993. Discoursal outcomes of small group work in an EFL classroom: A look at the interaction of non-native speakers. *Thames Valley University Working Papers in English Language Teaching* 2. 1–32.
- Foster, Pauline. 1998. A classroom perspective on the negotiation of meaning. *Applied Linguistics* 19(1). 1–23.
- Gass, Susan M. 2003. Input and interaction. In Catherine J. Doughty (ed.), *The handbook of second language acquisition*, 224–255. Oxford: Blackwell.
- Hendrickson, James M. 1978. Error correction in foreign language teaching: Recent theory, research, and practice. *The Modern Language Journal* 62(8). 387–398.
- Horgues, Céline & Sylwia Scheuer. 2014. “I understood you, but there was this pronunciation thing...”: L2 pronunciation feedback in English/French tandem interactions. *Research in Language* 12(2). 145–161.
- Horgues, Céline & Sylwia Scheuer. 2018. “What are you not understanding?” Managing communication breakdowns in native/non-native English/French tandem interactions. In Elena Babatsouli (ed.), *Crosslinguistic research in mono-*

- lingual and bilingual speech*, 129–146. Chania: Institute of Monolingual & Bilingual Speech.
- Horgues, Céline & Claire Tardieu. 2015. *Positionnement correctif dans les interactions orales en tandem anglais-français*. Paris: Association des chercheurs et enseignants didacticiens des langues étrangères ACEDLE. <https://halshs.archives-ouvertes.fr/halshs-01249735>.
- Horgues, Céline & Sylwia Scheuer. 2015. Why some things are better done in tandem? In Jose A. Mompeán & Jonás Fouz-González (eds.), *Investigating English pronunciation: Trends and directions*, 47–82. Basingstoke: Palgrave Macmillan.
- Jenkins, Jennifer. 2000. *The phonology of English as an international language*. Oxford: Oxford University Press.
- Kartchava, Eva. 2019. *Noticing oral corrective feedback in the second language classroom: Background and evidence*. Lanham: Lexington Books.
- Kartchava, Eva & Ahlem Ammar. 2014. The noticeability and effectiveness of corrective feedback in relation to target type. *Language Teaching Research* 18(4). 428–452.
- Keysar, Boaz. 2007. Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics* 4(1). 71–84.
- Li, Shaofeng. 2010. The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning* 60(2). 309–365.
- Loewen, Shawn. 2018. Focus on form versus focus on forms. In John I. Lontas (ed.), *The TESOL encyclopedia of English language teaching*, 2750–2755. Hoboken, NJ: John Wiley & Sons, Inc.
- Long, Michael H. 1991. Focus on form: A design feature in language teaching methodology. In Kees De Bot, Ralph Ginsberg & Claire Kramersch (eds.), *Foreign language research in cross-cultural perspective*, 39–52. Amsterdam: John Benjamins.
- Lyster, Roy & Leila Ranta. 1997. Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition* 19(1). 37–66.
- Lyster, Roy & Kazuya Saito. 2010. Oral feedback in classroom SLA. *Studies in Second Language Acquisition* 32(2). 265–302.
- Lyster, Roy, Kazuya Saito & Masatoshi Sato. 2013. Oral corrective feedback in second language classrooms. *Language Teaching* 46(1). 1–40.
- Mackey, Alison. 2006. Feedback, noticing and instructed second language learning. *Applied Linguistics* 27(3). 405–430.

8 Potential pitfalls of interpreting data from E-F tandem conversations

- Mackey, Alison, Susan M. Gass & Kim McDonough. 2000. How do learners perceive interactional feedback? *Studies in Second Language Acquisition* 22(4). 471–497.
- Mackey, Alison, Rhonda Oliver & Jennifer Leeman. 2003. Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning* 53(1). 35–66.
- Mauranen, Anna. 2006. Signaling and preventing misunderstanding in English as a lingua franca communication. *International Journal of the Sociology of Language* 177. 123–150.
- McKay, Sandra L. 2002. *Teaching English as an international language*. Oxford: Oxford University Press.
- Nassaji, Hossein & Eva Kartchava. 2017. *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*. New York/London: Routledge.
- Nassaji, Hossein & Eva Kartchava (eds.). in press. *The Cambridge handbook of corrective feedback in language learning and teaching*. Cambridge: Cambridge University Press.
- O'Dowd, Robert & Markus Ritter. 2006. Understanding and working with 'failed communication' in telecollaborative exchanges. *CALICO Journal* 23(3). 623–642.
- O'Rourke, Breffni. 2005. Form-focused interaction in online tandem learning. *CALICO Journal* 22(3). 433–466.
- Pallotti, Gabriele. 2009. CAF: Defining, refining and differentiating constructs. *Applied Linguistics* 30(4). 590–601.
- Saito, Kazuya. in press. Corrective feedback and the development of L2 pronunciation. In Hossein Nassaji & Eva Kartchava (eds.), *The Cambridge handbook of corrective feedback in language learning and teaching*. Cambridge: Cambridge University Press.
- Saito, Kazuya & Roy Lyster. 2012. Investigating the pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly* 46(2). 387–398.
- Sato, Masatoshi. 2017. Oral peer corrective feedback. In Hossein Nassaji & Eva Kartchava (eds.), *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*, 19–34. New York/London: Routledge.
- Sato, Masatoshi. in press. Metacognitive instruction for collaborative interaction: The process and product of self-regulated learning in the Chilean EFL context. In Craig Lambert & Rhonda Oliver (eds.), *Using tasks in diverse contexts*. Clevedon: Multilingual Matters.

- Sato, Masatoshi & Shawn Loewen. 2018. Metacognitive instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Language Learning* 68(2). 507–545.
- Sato, Masatoshi & Shawn Loewen. 2019. Methodological strengths, challenges, and joys of classroom-based quasi-experimental research: Metacognitive instruction and corrective feedback. In Robert M. DeKeyser & Goretti Prieto Botana (eds.), *Doing SLA research with implications for the classroom: reconciling methodological demands and pedagogical applicability*, 31–54. Amsterdam: John Benjamins.
- Sato, Masatoshi & Roy Lyster. 2012. Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition* 34(4). 591–626.
- Scheuer, Sylwia & Céline Horgues. 2020. Corrective feedback in English/French spoken tandem interactions. In Claire Tardieu & Céline Horgues (eds.), *Redefining tandem language and culture learning in higher education*, 147–160. New York, NY: Routledge.
- Sheen, Younghee. 2006. Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research* 10(4). 361–392.
- Sheen, Younghee & Rod Ellis. 2011. Corrective feedback in language teaching. In Eli Hinkel (ed.), *The handbook of research in second language teaching and learning*, 593–610. New York/London: Routledge.
- Sippel, Lieselotte & Carrie N. Jackson. 2015. Teacher vs. peer oral corrective feedback in the German language classroom. *Foreign Language Annals* 48(4). 688–705.
- Smith, Larry E. & Cecil L. Nelson. 1985. International intelligibility of English: Directions and resources. *World Englishes* 4(3). 333–342.
- Storch, Neomy. 1998. A classroom-based study: Insights from a collaborative text reconstruction task. *ELT Journal* 52(4). 119–158.
- Storch, Neomy. 2002. Patterns of interaction in ESL pair work. *Language Learning* 52(1). 291–300.
- Storch, Neomy & Ali Aldosari. 2012. Pairing learners in pair work activity. *Language Teaching Research* 17(1). 31–48.
- Swain, Merrill & Sharon Lapkin. 1995. Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics* 16(3). 371–391.
- Swain, Merrill & Sharon Lapkin. 1998. Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal* 82(3). 320–337.

8 *Potential pitfalls of interpreting data from E-F tandem conversations*

- Trudgill, Peter. 2005. Native-speaker segmental phonological models and the English Lingua Franca Core. In Katarzyna Dziubalska-Kořaczyk & Joanna Przedlacka (eds.), *English pronunciation models: A changing scene*, 77–98. Frankfurt am Main: Peter Lang.
- Varonis, Evangeline M. & Susan Gass. 1985. Miscommunication in native/nonnative conversation. *Language in Society* 14(3). 327–343.
- Ware, Paige D. & Robert O'Dowd. 2008. Peer feedback on language form in telecollaboration. *Language Learning & Technology* 12(1). 43–63.