

## Chapter 7

# Transcribing interlanguage: The case of verb-final [e] in L2 French

Pascale Leclercq

Université Paul Valéry Montpellier 3

This chapter aims at shedding some light on the place of transcription in the data interpretation process. More specifically, it focuses on the example of verb-final [e] in oral second-language French, which causes interpretation problems when context does not provide disambiguation cues. Through an analysis of three studies on this phenomenon (Herschensohn 2001; Prévost 2007b,a; Granget 2015), displaying a variety of theoretical frameworks (generative versus functional) and transcription options (written- and spoken-centric approaches), I show that transcription choices, whether made intuitively or in a theory-constrained manner, are often problematic as they entail an early categorization of data, even before data coding and analysis, thereby introducing an interpretive bias (Mondada 2007). Finally, I draw conclusions and offer suggestions regarding best transcription practices.

**Keywords:** Transcription, data interpretation, L2 French, verbal morphology, interlanguage

## 1 Introduction

This chapter stems from the author's questions and doubts while engaging in corpus-based second language acquisition (SLA) research, and more specifically when facing oral learner production that is ambiguous and requires the researcher to make conscious transcription decisions. I was particularly puzzled by the way some English-speaking learners of French use verb-final [e] as a generic verb ending, in spite of its highly polysemous value (e.g., *tomber* 'fall' infinitive, *tombait* 'was falling' imperfect, *tombé* 'fallen' past participle, etc.) and wondered how to transcribe and interpret such forms. This phenomenon has been pointed out regularly over the last decades in the literature on the acquisition of morphosyntax



and has been studied from a variety of theoretical perspectives (Myles et al. 1998; Herschensohn 2001; Bartning & Schlyter 2004), making it an interesting case for a study on the links between interlanguage description and interpretation (see for example the negation phenomenon, Ortega 2014). The description of learner language is indeed a complicated business, which involves, to a certain degree, inferring what learners want to say in a given context. How do researchers deal with this when transcribing raw data? How do they make such inferences? To what extent are their choices theory-constrained? This paper aims at shedding some light on these issues.

The concept of interlanguage (Selinker 1972; Han & Tarone 2014; Pallotti 2017) is key in SLA, as it reflects that a learner's language is a system in itself, and one which can vary in time and in contexts of use according to a number of variables (e.g., length and type of instruction, time spent in the target language country, motivation, aptitude, socioeconomic background, linguistic context, interlocutor). It depicts the language of learners as a dynamic unstable system, influenced by the patterns of the source language and/or other known languages, stabilizing at times, and sometimes subject to attrition or fossilization. Interlanguage development (notably speed and level of achievement) is constrained at the level of the individual, yet many researchers believe that there are shared itineraries (e.g., Bartning & Schlyter 2004), although this by no means implies that all developmental paths are the same. Researchers endeavouring to identify the dynamics of learning a language are faced with crucial methodological choices, regarding study design, data transcription and theoretical framework (Mackey & Gass 2012; Revesz 2012). In particular, they have to pay close attention to individual performance so as to find patterns, which may point to some general, and possibly universal, learning processes. Consequently, following Selinker's (1972) admonition to describe interlanguage before engaging in an explanatory process, SLA researchers have been working to map as accurately as possible the way learners of a new language develop their oral and written skills, whether in production or in comprehension (Ortega 2014). Oral production offers privileged access to the processes learners are engaged in when they utter messages in a foreign language, whatever their proficiency level: The repetitions, hesitations, and reformulations that are typical of the oral modality may tell the researcher whether the learners are able to plan and structure their discourse and utterances, through more or less automatized access to the second-language (L2) lexicon and grammar, and whether they are able to monitor their speech for errors (Segalowitz 2010: 47 cited by Hilton 2014: 29; Kormos 2006).

This brings us to a major issue facing the SLA researcher: The interpretation of oral interlanguage data, a process in which transcription plays a major part. As acknowledged by Ochs (1979) in her pioneering article, and much later by Mondada (2000; 2002; 2007), transcription is a theory-laden interpretive procedure, which incorporates the researcher's theoretical assumptions in the way oral phenomena should be represented and converted to written form.

Although transcription procedures are at the very heart of research on spoken language, the transcription process itself has seldom been explored, and even less so from an SLA perspective. Yet, when transcribing learner oral production, the researcher has to make a series of strategic decisions regarding how to interpret ambiguous forms, such as, for example, verb-final [e] in L2 French, which can stand for homophonous infinitive (*tomber*), imperfect (*tombais, tombait*), past participle (*tombé, tombés*), or passé simple (*tombai*) marks in standard French, or it might even stand for something else in the learner's interlanguage, such as simple present (Granget 2015). How should such forms be represented in transcription? Should researchers use orthographic conventions and take a decision based on context and their knowledge of the target-language norms (for example, a past adverbial may lead researchers to adopt the interpretation of imperfect or *passé composé*), or should they leave the interpretation open and use phonetic transcription (MacWhinney 2000: 19; Saturno 2015)?

Against this backdrop, I purport to explore the way verb-final [e] in L2 French has been transcribed and analysed in two theoretical perspectives, with the aim of contributing to the current discussion on interlanguage description and interpretation (Ortega 2014), as well as offering some methodological reflections on data transcription.

Keeping in mind MacWhinney's (2000: 19) warning that "perhaps the greatest danger facing the transcriber is the tendency to treat spoken language as if it were written language", I introduce the verb-[e] (henceforth V-[e]) transcription problem, and I reflect on the task the transcriber faces as well as on the use of the French writing system to transcribe oral learner data. Then I present the transcription, coding and interpretation choices made by three researchers (Herschensohn 2001; Prévost 2007b,a; Granget 2015) regarding the use of verb-final homophonous [e] in L2 French. These three studies were selected as they feature an in-depth analysis of the [e] phenomenon, while offering different theoretical perspectives and transcription strategies (orthographic or phonetic). Based on these three approaches, and following Ortega (2014), I conclude by discussing the link between the choice of theoretical perspective and the description of linguistic data through the transcription process and providing a few guidelines for transcribers.

## 2 The verb-final [e] problem in transcription

My first encounter with verb-final [e] took place when I was preparing a paper on how learners' way of referring to time and space in narrative discourse developed over the course of L2 acquisition and whether the data corroborated the Aspect Hypothesis (Leclercq 2011).

Through a focus on motion predicates, in combination with tense and aspect markers, and within a crosslinguistic perspective, I wanted to provide a characterization of elementary, intermediate and advanced proficiency levels. The main research question was the following: How do learners' ways of referring to time and space develop over proficiency levels? I hypothesized that the selection of motion predicates was closely linked to morphological aspect marking (aspect hypothesis, Andersen & Shirai 1994; Robison 1995, also see Rohde 1996 for an overview), and that ambiguous forms as regards tense/aspect marking in English would decrease over proficiency levels.

The experimental design consisted of the administration of a biographical questionnaire, which included information regarding the language-learning history of the participants, and the completion of the Horse Story, an oral picture-based story-telling task developed within the Langacross project (a Franco-German research project funded by the French National Research Agency (ANR) and supervised by Maya Hickmann). The stimulus featured five pictures in which three entities (a horse, a cow and a bird) were localized with reference to a meadow and a fence. It triggered the retelling of motion events (running, jumping, falling, flying) (see Appendix). It was initially used to study the acquisition of spatial reference in first-language (L1) and L2 French, Chinese, German and English (Hendriks 1998; Hickmann et al. 1998). The retellings were recorded and later transcribed by a trained researcher using CLAN conventions and the recommendations provided by the Langacross team in their coding manual (Hickmann et al. 2011), after which I checked them myself. In accordance with those recommendations, a @G line was inserted in the transcriptions to indicate the correspondence between the learners' utterances and the picture from the stimulus. For the purpose of the present chapter, the target verb-final [e] forms are bolded in the following transcription.

I now examine example (1), which presents a retelling by an intermediate-level English learner of French, Mag, including verb-final [e] forms. I highlighted those forms in the transcription.

- (1) Horse Story, 2011 transcription, English L1, French L2 Mag, intermediate.
- @G: pict\_1  
 \*SUJ: *dans une première c' est un cheval* . 'in a first it is a horse'  
 \*SUJ: *qui court dans une prairie* . 'which runs in a meadow'  
 @G: pict\_2  
 \*SUJ: *mm c' est mm une vache dans le prochain image* . 'it is a cow in the next picture'  
 \*SUJ: *séparée par une mm un pré mm avec un oiseau* . 'separated by a field with a bird'  
 @G: pict\_3  
 \*SUJ: *mm alors dans le troisième le cheval essayE de monter le* . 'so in the third the horse try to climb the'  
 \*SUJ: *comment on dit* . 'how do you say'  
 \*EXP: *la barrière* . 'the barrier'  
 \*SUJ: *la barrière oui* . 'the barrier yes'  
 \*SUJ: *mm mais le barrière se casse* . 'but the barrier breaks'  
 @G: pict\_4  
 \*SUJ: *dans le quatrième est le cheval tombE* . 'in the fourth is the horse fall'  
 @G: pict\_5  
 \*SUJ: *et dans xxx le vache et l' oiseau euh aidE le cheval avec une boîte*  
 <de> [/] # *de aide et # une bandage* . 'and in xxx the cow and the bird help the horse with a box of help and a bandage'

What the learner produces is several instances of verbal forms with a verb-final [e] sound, which could either stand for an infinitive, or a past participle, or an inflected form (imperfect *aidait* 'helped' for example), or which could constitute a base form in the learner's interlanguage. What is interesting is that this particular learner used the [e] ending on a regular basis, but also used targetlike verb forms like present tense *se casse* 'breaks down'. The researcher, when facing forms that are ambiguous, has to make decisions regarding transcription and interpretation of data.

Example (1) shows that I chose to use regular orthographic spelling when phonological forms seemed to match the tense/aspect/person agreement rules, as is the case with *séparée* 'separated' on line 4 of the transcription, and the V-E symbol only for the forms which I identified as potentially ambiguous, either due to a position in the sentence that might be interpreted as necessitating an inflected verb (either present tense *aide* with a mute final <e> or passé composé *a aidé* 'helped'), or due to unusual word order (*est le cheval & tombE* 'is

the horse fall-ED’). In this project I did not even consider that forms like *séparée* ‘separated’ could actually be non-targetlike in the mind of the learner, and based on our knowledge of French graphonemics, the transcriber and I assumed that learners had produced target forms (a fairly naïve and controversial position). These intuitive transcription choices could be qualified, in the words of Ortega (2014), as pre-theoretical. Mondada (2000: 3) nevertheless points out that such intuitive choices are highly problematic when it comes to data categorisation and interpretation.

I will now try to shed some light on this phenomenon by considering the transcription process itself.

### 3 Transcribing as a situated practice and a cognitive challenge

Although it is a fundamentally theoretical enterprise (Ochs 1979), and a crucial part of the research process, transcription is a grey zone in most studies. This is reflected in the scarcity of research on the rationale and consequences of transcribers’ choices (Mondada 2002: 46). In a series of papers using Ochs’ study as a starting point and a framework for the analysis of the activity of transcription from an epistemological perspective, Mondada (2000; 2002; 2007) proposes an in-depth description of the transcription process. Bearing her analyses in mind, I discuss the practical and cognitive challenges awaiting the transcriber when it comes to the transcription of verb-final [e] in L2 learner data.

#### 3.1 What is transcription?

First, Mondada (2007: 810) refers to transcribing as a “situated practice”, observing that it is “embedded within a series of research practices: data production, digitalization and compression, anonymization, storage and filing, representation and annotation, analysis, and so on [...]. These practices configure and more radically ‘fabricate’ what we consider as ‘data.’” She acknowledges the fact that transcripts on their own are not data since they cannot be autonomized from recordings (which constitute primary data). Transcripts are considered “secondary products of representation and annotation practices” (Mondada 2007: 810–811). Transcripts and recordings complement each other, particularly when data transcription software such as CLAN is used, as they enable the linking of transcription with the original audio or video recording. While transcripts enable researchers to focus on detail for analysis, recordings provide the possibility to listen again.

Mondada therefore acknowledges the evolutive nature of transcripts: The transcriber can endlessly check, revise, reformat, for a specific analysis or for editorial purposes (p. 810). Another inherent feature of transcribing is variability (both within and in-between transcriptions, as illustrated in the different treatments of *séparée* and *essayE* in (1)).

On a more abstract level, Mondada (2000) describes the transcription process as an exploitation of writing resources in order to create a representation of oral discourse based on operations of filtering "noises" (phenomena deemed non-meaningful by the transcriber), and of homogenisation through the use of systematic conventions (this latter point is particularly well exemplified in the CLAN and CHAT manuals, MacWhinney 2000). Mondada (2000) observes that the passage from oral to written form has consequences on the interpretability of the spoken language. On the one hand, the transcription appears as a structured account of oral speech, facilitating visual perception. On the other hand, having the possibility to listen again and again to the same recording provokes what Mondada refers to as a magnifier effect: The researcher can focus on a phenomenon which is ephemeral in real-time and might have passed unnoticed during normal conversation.

In line with Ochs (1979), Jefferson (1996) and Saturno (2015), Mondada (2000) indicates that transcribing is an inherently interpretive activity. She also highlights what she calls the circularity problem: Numerous interpretations of phenomena are incorporated a priori in the transcription, in spite of their being the purported aim of the a posteriori analysis of this transcription. In other words, the transcription choices made by the researchers already contain their interpretative choices, making the whole research enterprise dubious.

Mondada (2000: 8) also describes the transcriber's job as "isolating, cutting out, identifying, making identifiable the recorded forms in a clear written form."<sup>1</sup> Some notation systems enable the highlighting of indeterminate forms. Among those systems, she mentions the use of International Phonetic Alphabet (IPA) as opposed to orthographic representation. She claims that using IPA suggests a form is deemed non-identifiable by the researcher, while an orthographic rendering indicates that the researcher has already categorized the form. The use of phonetic versus orthographic conventions shows that transcribed forms can be categorised as "more or less comprehensible (transparent or opaque), more or less standard (according to their distance with the standard), [...] a social category, etc." (Mondada 2000: 8, my translation). While complete IPA transcriptions are often deemed impractical (and are rarely seen in SLA projects), Mondada points

---

<sup>1</sup>My translation.

out that the punctual use of IPA allows us to avoid selecting an orthographic form, and therefore a specific language (a crucial point in interlanguage studies), but also to visually detach the transcribed content and consequently highlight it. I will refer to the orthographic option as a written-centric approach, and to the punctual IPA option as a spoken-centric approach. Although the spoken-centric option appears as more careful, as it leaves data interpretation open, it is not always chosen by researchers. Why is that? An interesting hypothesis is that the written code deeply influences literate speakers' representation and categorization of language units (Jaffré 2006, see below).

Keeping Mondada's reflections in mind, I now focus on the choice of orthographic transcription and what it might imply for the transcriber, particularly how phonological and graphical representations may interact and be treated by the transcriber. I use the verb-final [e] phenomenon as a basis for the analysis.

### **3.2 Cognitive aspects of the learner and transcriber's task: Making the most of homophony in spoken French**

While most practitioners and researchers agree that verb morphology, including verb-final [e] in French, is a major source of spelling confusion, whether for French children learning how to write (Brissaud & Sardon 1999; David et al. 2006; Fayol & Pacton 2006), for adults who master the orthographic system and use it daily, or for L2 learners in their oral and written productions (David et al. 2006; Brissaud et al. 2006; Prévost 2007b; Granget 2015), some researchers have started focusing on the impact of the orthographic experience on phonological awareness (Bassetti et al. 2015; Nimz & Khattab 2019). According to Detey & Nespoulous (2008: 68), several studies suggest that "orthographic representations might play a role in speech perception [...], at least through bidirectional activations between graphemes and phonemes [...]. Hearing a lexical unit might activate an orthographic representation, which might in turn influence phonological judgements." As a consequence, it is quite possible that experts' interpretation of oral learner speech might be largely influenced, directly or indirectly, by spelling knowledge. The transcriber, whether a native speaker or an expert user of the target language, as is the case in (1), often tackles the data with a number of assumptions about what the learner knows about the language, including morphographic representations.



### 3.3 Description of verb-final [e] phenomenon in relation with the French spelling system

According to Jaffré (2006: 25), a spelling system is not a mere tool used for the sake of written communication: Centuries of usage have fostered tight links between spelling forms and their users, whose perceptive abilities they have shaped and constructed. He therefore argues that beyond the strict communicative need to disambiguate homophones, their very orthographic differentiation has progressively conditioned the cognitive representations of literate users. In that context, the written code can be considered as an autonomous linguistic representation, capable of exerting an influence on spoken language, as shown by the orthographic distinction of homophones in French. Of course, such an influence is only possible in a society where written communication forms part of the essentials of daily life, and where literacy is a basic skill most citizens master, as is the case in 21<sup>st</sup> century France.

According to Jaffré (2006), spelling systems have two main objectives. First and foremost, they aim at representing a given language phonographically, but they also have a semiological agenda: They aim at providing a visual representation of language, with tools that enable the disambiguation of spoken forms, for them to be readable and interpretable. In that regard, the use of orthographic spelling for L2 transcriptions compels transcribers to make choices informed by their knowledge of the target language grammar. Additionally, although languages such as French and English use an alphabetic system, their spelling systems often depart from a one-to-one sound/letter relation. In that context, heterography appears as a tool to tease apart homophones.

The French morphological verbal system, and more specifically final [e], provides an interesting case as regards homophony and heterography, and one which poses a particular challenge to the learner. As noted by Brissaud et al. (2006), a form such as /truve/ can be spelt in ten different ways (*trouver, trouvé, trouvés, trouvée, trouvées, trouvai, trouvais, trouvait, trouvaient, trouvez*), making the native writer and the L2 learner's task a daunting one. Under the umbrella of verb-final [e] forms, there are two types of combining morphemes: Mute endings (-e, -s, -t, -ent), referring to person, gender and number; and tense, aspect and morphemes (among others, -ai, -é, -er, respectively referring to imperfect, perfect and infinitive). Context plays a major role in the selection of the appropriate lexical item, and as a result, the appropriate spelling. However, it is often not sufficient to disambiguate homophonous conjugated verbs, as is the case between infinitive *jouer* and past participle *joué*, which are a frequent source of spelling mistakes, even among French native speakers. Scriptors then have

to rely on their knowledge of grammatical rules to select the right target form (David et al. 2006; Brissaud et al. 2006). Regarding verb-final [e], French scribes have usually received explicit instruction at school as to verb-agreement rules (subject-verb agreement and tense/aspect endings), grammatical functions and categories (such as infinitive and past participle), and linguistic manipulations to disambiguate homophonous forms (for example, if /ale/ can be replaced by /partir/, it has to be an infinitive form spelt *aller*.) They may also rely on their memorization of frequent co-occurring units (for example, *pour* is usually followed by an infinitive form, see Brissaud et al. 2006: 78). In a written-centric approach, the transcriber's interpretation of ambiguous oral forms consequently requires an analysis of context, and it taps into their phraseological knowledge as well as their knowledge of the target language orthographic rules, in order to apply the relevant contextual, categorical or morphological rules.

### 3.4 A challenge for the researcher

When transcribing learner data, using a written-centric approach is a leap of faith, as there is no guarantee at all that the learners master the system in the same way as a native speaker. Transcribing then becomes a game of inference: The researcher tries to infer what the learner had in mind when uttering a given phrase, and makes hypotheses regarding their choices, possibly in accordance with the selected theoretical framework of analysis. Transcription problems are consequently inherent to the fact that transcribers often assume that learners share their knowledge of the target language, including its writing system. Flavell (1977, cited by Nickerson 1999: 739) observes that “we are usually unable to turn our own viewpoint off completely when trying to infer the other's, and it usually continues to ring in our ears while we try to decode the other's. It may take considerable skill and effort to represent another's point of view accurately through this kind of noise, and the possibility of egocentric distortion is ever present.” It is therefore important for the researcher to keep in mind that they do not know the extent of the L2 learner's knowledge of the target language and culture. Shared knowledge, and an understanding of what the interlocutor knows, is at the heart of the communication process (Nickerson 1999; Keysar et al. 2003). We assume our interlocutors share basic communication principles. However, as Nickerson (1999) puts it, “overimputation” of one's knowledge (i.e., attributing learners' knowledge about the target language that they do not necessarily possess), or lack of ability to adopt another perspective than one's own, can cause communication difficulties and lead to an incorrect interpretation of the interlocutor's message. The researcher has to be careful about “overimputation” and

has to remember that nothing can be taken for granted in the realm of L2 acquisition. We cannot assume that an L2 learner thinks in the exact same way as a native speaker, nor that they use verbal forms with the same degree of mastery of form/function relations. When learners use verb-final [e] forms, do they try (with mitigated success) to retrieve targetlike morphology, or do they create a new ending to compensate for their lack of procedural knowledge? When in doubt, using a spoken-centric approach (i.e., using IPA to transcribe ambiguous forms) might be a good option, as it does not entail early categorisation of the data.

Keeping this in mind, we will examine how verb-final [e] was transcribed and analysed in three different studies (Herschensohn 2001; Prévost 2007a,b,<sup>2</sup> Granget 2015), displaying a variety of theoretical frameworks (generative vs. functional) and transcription options (written and spoken-centric approaches), so as to shed light on the link between data transcription and interpretation, and offer reflections regarding best transcription practices.

## 4 Making theoretically-informed transcription and interpretation choices

As stated by Ortega (2014: 186), “the formal linguistic study of L2 development puts theory first and is driven by the quest to understand the role that Universal Grammar or abstract linguistic knowledge plays in the acquisition of human language across the life span.” Generative researchers therefore favour a top-down approach, with overarching research questions focusing on finding proof in favour of (or against) the theoretical constructs under scrutiny through data analysis (Lardière 2012). Data are used to inform or contradict theoretical premises. This is the case with L2 morphosyntax and particularly with the development of inflection (Herschensohn 2001), which have been largely explored within a generative, theory-constrained framework (among others, Prévost & White 2000; Prévost 2007a; Herschensohn 2001; Hawkins 2004). On the other end of the theoretical spectrum, functionalist researchers (see Lenart & Perdue 2004) take a rather bottom-up, data-driven approach: Based on learner discourse, they try to account for the way learners acquire and make use of formal linguistic levels of organisation (morphology, morphosyntax, lexicon) in a given context of use. In the words of Klein & Perdue (1989) (cited by Lenart & Perdue 2004:

---

<sup>2</sup>Prévost (2007a) and Prévost (2007b) present complementary information about the study under consideration in this chapter.

85), researchers have to solve “the learner’s problem of arranging words” to produce a contextually meaningful message. Functionalists have also provided detailed accounts of the verb-final [e] in French L2, through longitudinal studies (for example, studies based on the ESF project, such as Noyau et al. 1995; Véronique 2004), or cross-sectional studies (Bartning & Schlyter 2004; Granget 2015). I compare theory-constrained approaches and data-driven approaches to see what each contributes to the debate on how to interpret, and thus transcribe, oral interlanguage productions, as regards the analysis of verb-final [e]. This section was also largely inspired by the work of Granget (2015), who paved the way for the following analysis by citing the work of Herschensohn (2001) and Prévost (2007b).

#### **4.1 Theory-constrained and written-centric approaches: Defective or missing surface inflection hypothesis**

The verb-final [e] phenomenon is presented by Herschensohn (2001) as part of a wider debate on the development of inflection in learner grammars and what it reveals about access of L2 users to Universal Grammar. Within a generative framework of analysis, Herschensohn seeks to determine whether the empirical data support theoretical claims regarding the reason why L2 learners’ verbal inflections are so often defective (i.e., not targetlike) especially at intermediate levels. She re-examines the relationship between the acquisition of morphology and functional categories: Some researchers claim that morphology and syntax are developed conjointly in the L2 grammar (co-dependence, Eubank 1993; Vainikka & Young-Scholten 1996; 1998a,b cited by Herschensohn 2001), in the same way as in L1 grammar, while others reject such a link and propose that morphology and syntax develop independently (Schwartz & Sprouse 1996; Lardière 1998 cited by Herschensohn 2001). In this view, a lack of morphological marks is attributed to processing difficulties. L2 learners may display evidence of syntactic competence, but deficient morphological production because “under the Missing Surface Inflection Hypothesis, the L2er has a grammar with complete functional projections but incomplete morphology” (Herschensohn 2001: 280).

Within this framework, Herschensohn (2001) argues that inflectional deficits (i.e., non-finite verbs, or other morphological errors) support neither the co-dependence nor the independence hypothesis. She claims that “the French data rather indicate that deficiencies in morphological mapping, not defective syntax (functional categories) are the cause of L2 failed inflection” (p. 273). What is very interesting in this paper, and which provides one of the main reasons to include it in the current analysis, is that the author explicitly motivates her choice of

using the infinitival form in the transcription, making her approach a typically written-centric one. Tapping into the abundant literature on the acquisition of morphosyntax, she situated her transcription choice of attributing an infinitive value to verb-final [e] forms from the very onset of the paper: “the second language (L2) use of *infinitives in contexts of obligatory tense* – amply documented and discussed in the literature [...] – is of theoretical interest [...]” (p. 273) (emphasis mine).

Verb-final [e] is examined in this study based on data collected from two stay-abroad participants, Chloe and Emma, through a series of interviews that took place before, at midpoint and after their six-month period in France. The interviews include discussion of topics in the present tense, and questions that refer to past and future actions.

The data were transcribed by the author herself and checked by a French phonologist. It yielded “several hundred tokens of verb morphology”, among which “a number of verb errors are transcribed as infinitives, although the infinitival form is homophonous with the past participle and the *vous* form of the present in most cases” (Herschensohn 2001: 285–286). Herschensohn: 286 explains the rationale behind her choice of transcription for these forms: “In the cases where the form is transcribed as an *-er* infinitive, the context clearly excludes the possibility that the inflection is a past participle.” She provides the two examples in (2) and (3).

- (2) *A quelle heure est-ce que le musée ouvrir?* (p. 286)

At what time Q the museum open

- (3) *A quelle heure est-ce que le musée fermer?* (p. 286)

At what time Q the museum close

She explains this choice as coming from her interpretation of such interlanguage forms as “the regularization of the irregular infinitives *ouvrir* (‘to open’)” rather than as the imperfect *ouvrait* (‘it was opening’)” ((Herschensohn 2001: 286)). She notes that when she transcribed the [e] form as an infinitive, the other options (*vous* form *ouvrez*, past participle, imperfect form *ouvrait*) were excluded by contextual information. The author also explicitly rejects an interpretation of those forms as occurrences of the imperfect (*ouvrait, fermait*), as she claims that “the use of any tense other than present would be inappropriate” (p. 286). However, what does “inappropriate” mean in the context of a learner production? In spite of the author’s claim, it seems difficult to rule out the imperfect interpretation, which would be grammatically correct, though less expected than simple present in the cited context. This might well be a case of “overimputation” on

the part of the researcher (Nickerson 1999), who might have been influenced by her knowledge of French written rules, as proposed by Jaffré (2006).

In sum, Herschensohn (2001) showcases how theory constrains transcription and interpretation choices: The author's decision to make theoretically informed choices regarding the spelling, and therefore the grammatical status, of verb-final [e] forms, entails that these morphological forms are ascribed a specific (potentially erroneous) value even before the data description stage, thereby constraining interpretation. This is what Mondada (2000: 2) refers to as the circularity problem: The researcher makes theory-informed data interpretation decisions at the transcription level and then analyses the transcribed phenomenon in the light of the same theoretical framework.

Another example of theoretically informed labelling is provided in Prévost (2007b), in which he refers to verb-final [e] forms as "default forms" in his study on the influence of the source language verbal paradigm on morphological variability in L2 French. He describes L2 French infinitives as recurring erroneous forms used in contexts where other forms are expected in the target language ("Ah je voyager >à> [/] à des Etats <U> [/] Unis," Prévost 2007b: 50). He notices that these forms are used in a similar way by child and adult learners, of various mother tongues (English, Turkish or Chinese), and at different proficiency levels (beginners to advanced). He also points out the fact that these forms are not the only ones that are used by learners and that they can co-exist with targetlike forms, often within the same sentence. He calls this phenomenon morphological variability. Like Herschensohn (2001), he situates his research within a generative framework, and asks whether morphological variability reflects some sort of deficiency of interlanguage grammars, especially as regards functional categories such as Infl(ection).

Prévost (2007a) analyses verbal errors produced by 21 Anglophone learners of French, at four proficiency levels (beginner to upper intermediate 3). He observed that learners produced infinitives in contexts where an inflected verb was expected, such as after a lexical subject, but they also used inflected forms in contexts where an uninflected form was expected, for example after a preposition, an auxiliary or another verb. Prévost (2007a: 360) explains his transcription choices in the following way: "[A]n infinitival verb was considered nonfinite unless evidence of the contrary existed. In other words, verbs ending in [e], which is ambiguous between the infinitival marker *-er*, the past participle marker *-é*, and the second-person plural ending *-ez*, were categorized as nonfinite unless they appeared with the second-person plural pronoun *vous* 'you'." Just like Herschensohn (2001), Prévost adopts a written-centric approach and does not envisage the

possibility of an imperfect meaning or of another undetermined value. What is more, his explanation perfectly illustrates the fact that transcription choices imply a categorisation of data (and hence, constrain data analysis) (Mondada 2007).

The examples provided in the paper show that at the level of transcription, the author chose the *-er* infinitive ending whenever an ambiguous form occurred:

- (4) a. *Je parler avec administration* (Ann, G1)  
I speak.INF with administration
- b. *Ils visiter moi.* (Mike, G2)  
they visit.INF me
- c. *On aller au centre d'achats* (Jill, G3)  
we go.INF to mall
- d. *Il retourner à la maison* (Sandra, G4) (p. 363)  
he return.INF to the house

He justifies his choice by stating that similar constructions also occur with verbs of the second (*ouvrir*) or third (*boire*) group, as in (5a, c) and are also found in his data with negation as in (6):

- (5) a. *Il boire* (Edward, G1) (p. 361)  
he drink.INF
- b. *Il prendre des vêtements* (John, G4) (p. 361)  
he take.INF some clothes
- c. *... quand ils ouvrir les cadeaux* (Kate, G3) (p.364)  
when they open.INF the presents
- (6) *Je ne expliquer pas* (Mark, G1) (p. 365)  
I NEG explain.INF NEG
- (7) a. *J' ai difficile demande des questions* (Ann, G1)  
I have.1s difficult ask.s some questions
- b. *Comment tu vas arrive à mon travail ?* (Jen, G2)  
how you go.s to my work
- c. *... qui uh j' ai rencontre à Nouvelle Ecosse* (Jill, G3)" (p.369)  
who I have.1s meet.s in Nova Scotia

Examples from Prévost (2007a), presented in Prévost (2007b: 61)

Just like Herschensohn, Prévost (2007a) refers to those forms as errors. What is very interesting in Prévost (2007a) is that he observes and describes substantial inter-individual variation in the use of uninflected forms where an inflected form is expected, and in the use of inflected forms where uninflected forms are expected as in (7). Based on the fact that the learners in this study are Anglophones and that their L1 does not possess a particular marker for infinitive, he interprets their choice of inflected forms where uninflected ones are expected as some sort of base uninflected form for the learners. He also very perceptively observes that many errors share an ambiguous phonological form (*je/tu/il/elle/on/il-s/elles chant[t]* ‘I/you/he/she/it/they sing’), and that they could very well be uninflected forms in the mind of learners. Although his overarching research questions are driven by the generative-approach agenda, and in spite of transcription choices that are highly constraining as they constitute a pre-analysis and pre-interpretation of the data, Prévost (2007a) proposes a careful description of the data under scrutiny and is wary of overinterpretation. We will now turn to a more strictly data-driven approach with the analysis proposed by Granget (2015).

## 4.2 Data-driven and spoken-centric approaches

Granget (2015) investigates the acquisition of the French present tense, also called “présent simple”, and more particularly the emergence of the morphological expression of this tense in L2 French, and she adopts a radically different approach. She used a database of 36 oral picture story retellings (Loch Ness stimulus from the French Learner Language Oral Corpora (FLLOC) project, Marsden et al. 2002) by Anglophone teenagers learning French, at three different institutional levels (3 to 5 years of instruction, 12 learners per level).

Granget (2015) described the database under investigation, and emphasised its interactional nature: During the task the learners interacted with the interviewer, who provided either positive feedback (*très bien*, ‘very good’, *c’est bien* ‘it’s good’ etc.) to answer the learners’ vocabulary requests or helped them move the narrative forward (*En bateau très bien, qu’est-ce qu’elle fait?* ‘In a boat very good, what is she doing?’). She then explained her transcription choices: Although the FLLOC recordings were transcribed following CLAN procedures, Granget modified the original transcriptions where ambiguous forms occurred. More specifically, she “demorphologized” ambiguous verbal forms and transcribed them phonetically in order to question the learners’ morphological representation of these forms. For example, when the learner asks the interviewer “what’s fishing?” (p. 123), the reply [pɛʃ] is transcribed phonetically because Granget did not suppose a priori that the learner used an inflected form, as the use of *pêche* ‘to fish, simple



present, third-person singular’ or *pêchent* ‘to fish, simple present, third-person plural’ would indicate. Non-ambiguous irregular verb forms such as *sont* ‘be third-person plural’, *est* ‘be third-person singular’, *fait* ‘do third-person singular’ are nevertheless transcribed using orthographic conventions.

Granget (2015) sought to answer the following research question: To what extent can we consider that L2 learners’ verbal forms include morphological markers? For example, in (8), is [sorte] ‘go out’ an inflected form (base form [sort] + *e* morpheme)? In (9), [di] sounds like a native-like inflected form (*dit*, ‘says’) but is this the case from the learner’s view point?

(8) *le mère et le petit garçon* [sorte] *les villes*  
 ‘the mother and the little boy go\_out the city’

(9) *les enfants et le grand-mère* [di] *au revoir*  
 ‘the children and the grandmother say bye’

She performed a qualitative analysis using the Finiteness framework (Klein 2006). According to Perdue et al. (2002: 853) (see also Klein 2006 and Gretsche & Perdue 2007), finiteness “is usually associated with the morphosyntactic categories of person and tense”. However, Perdue et al. distinguish between the finiteness concept and the markers used to express it in the world’s languages. Klein (1994) relates finiteness to assertion (i.e., “the speaker’s making a claim about a time span” Perdue et al. 2002: 853). This implies temporal and aspectual anchoring (Klein 1994). The authors distinguish two types of finiteness:

- S(emantic) finiteness is a property of the whole utterance and is often marked by discursive or lexical means.
- M(orphological) finiteness denotes finite verbal morphology.

Within this framework, the learner’s tasks are the following: (a) noticing and acquiring the means that the target language provides for the expression of S-finiteness, and (b) finding out whether there are grammaticalised means to express (M) finiteness. Development from S (i.e., discursive or lexical means) to M (i.e., morphological means) finiteness denotes a progression towards native norm, as illustrated by these examples from the ESF project: (examples (10) and (11) come from Véronique 2004: 267, example (12) comes from Granget 2015: 114).

(10) Lexical and pragmatic means : [e] *la dame la cassette comme ça*. (Zarah)

(11) Lexical means : *moi [iparle] comme la cassette* (Zarah)

- (12) Morphological means : *Ma fille elle va déjà au lycée* (Alfonso)

In (10) and (11), Zarah describes a French class she attends. She relies on her interlocutor's capacity of inference, and on pragmatic means (in (10) she uses a gesture to show how the lady plays the cassette), while using lexicalised form [iparle]. In (12), Alfonso uses a targetlike verb form.

When coding, Granget (2015) was careful to use labels that do not predetermine the finite/non-finite status of the target elements: V-[e], V, Aux + V. Her analysis suggests that the same verb is often used within the same production with phonological variations that might reveal morphological variation ([rəgard]/[rəgarde]), an observation also made by Prévost (2007a,b). She tries to account for the distribution of such allomorphs in the corpus and envisages several possible explanations for this phenomenon. First of all, the presence of a final [e] morpheme in some but not all verbal forms makes it difficult to decide whether [e] can be interpreted as a flexional morpheme or whether forms such as ([rəgard]/[rəgarde]) both belong to the mental lexicon. She highlights the extreme difficulty in using target-language functional categories such as tense/aspect to describe learners' interlanguage. She then evokes the Aspect Hypothesis to account for the use of verb-final [e] in the data. According to this hypothesis, verb-final [e] would occur more frequently with predicates denoting bounded events. However, she observes that it is difficult to determine the lexical aspect of some of the verbal forms used by learners. In (13), it is difficult to assert whether [rəturne] is telic (taking the buoys out of the lake) or atelic (describing the state when the buoys are on the bank).

- (13) *le grand-mère [rəturne] les bouées de le lac*  
? 'the grandmother turn\_round the buoys of the lake'

Granget (2015: 132) therefore decides against the Aspect Hypothesis to account for the use of those forms and concludes that these forms should be treated as non-finite and non-analysed and that free variation is the rule in learner discourse.

- (14) A30: *l'enfance et (.) le le mère (.) euh [rəgarde] la monster dans la lac*  
ADR : *mmm*  
A30 : *euh (.) un [?] euh (.) un journaliste et touriste*  
ADR : *mmm*  
A30 : *euh [rəgard] la monster euh*  
A30 : *euh (.) maintenant euh l'enfance le enf les enfants [rəgarde] la tele*

Granget (2015) finally states that the data show high variability in the means of temporal anchoring and linking within and between narratives and advocates for the use of descriptive labels that avoid a pre-analysis assignment of linguistic category, in an effort not to overinterpret learner data. She suggests V-[e] forms can be interpreted as ‘verboid means of assertion’ (i.e., lexical forms with syntactic properties of finite verbs). As for whether these forms display morphological properties of inflected verbs or not, Granget (2015: 132) interprets them as “non-inflected and non-analysed verboid forms”.

Although these three studies focus on a common phenomenon (i.e., the use of verbal inflection in the oral production of L2 learners of French), the authors’ decisions regarding data transcription were all firmly anchored in their theoretical framework. Our comparison shows that decisions at the transcription level condition the description of a given phenomenon and its interpretation. In fact, the transcription stage is strongly dependent on theoretical assumptions. Depending on their overarching goals and on their theoretical framework, researchers may opt for transcription choices that reveal a pre-categorisation of the data (e.g., adopting the *-er* infinitive spelling for verb-final [e], a written-centric approach), or for earmarking ambiguous forms for future analysis, such as using the phonetic symbols instead of deciding on a specific spelling (spoken-centric approach), thus keeping all interpretive options open.

## 5 Conclusion

Through this chapter, I aimed to shed light on the transcription process in the context of research based on learner corpora, more specifically when oral production tasks are involved. Transcription indeed constitutes a crucial step in the constitution of oral corpora, as it shapes the data and makes them ready for subsequent analyses. The current study briefly described the cognitive aspects of transcription and focused on the methodological implications of transcription choices (written vs. spoken-centric approaches).

First, I tried to describe the difficult task facing the researcher when transcribing learner data, and more particularly in the case of ambiguous homophonous forms, such as verb-final [e] in French L2. Data transcribers often work on the assumption that learner language can be safely mapped onto the target written orthographic system and often rely on phonological and contextual cues provided by the learners to process and make sense out of ambiguous sounds. However, this entails a risk of “overimputation” on the part of the researcher (Nickerson 1999). This in turn creates methodological problems, as data descrip-

tion, analysis and interpretation are highly dependent on the transcription process itself (Mondada 2007). For example, transcribing ambiguous verb-final [e] as an infinitive *-er* form, as Herschensohn (2001) and Prévost (2007a,b) do, reveals pre-analysis choices that entail a labelling of those forms as errors, thus conditioning the subsequent analysis. This subsequently leads us to questioning the validity of analysing learner data in the light of native speakers' productions. In Dewaele's (2003) homage to Larry Selinker, one of the founding fathers of the interlanguage concept, he notices that in spite of the enormous success of this concept, linguists persist in comparing interlanguage with native speakers' systems and tend to analyse any deviation from the norm as a deficiency on the part of the learner. He then praises Cook's (2002) plea for learners to be considered as language users rather than as failed native speakers. We believe this is sound practice and we wish to encourage researchers to beware of overimputation or overinterpretation of learner data. When possible, we assume the best transcription practice is to adopt a spoken-centric approach and earmark the problematic forms through the use of IPA, without taking any decision regarding interpretation at the transcription level, as proposed and exemplified by Mondada (2002), Granget (2015), and Saturno (2015).

As for interpretation, I hope to have shown that it heavily depends on theoretical framework. It is also closely linked to the transcriber's intimate knowledge of the writing system and the assumption that such knowledge is shared by learners. However, we have no way of verifying this assumption from listening to a recorded production. Indeed, it remains to be understood why learners of French use verb-final [e]. None of the three papers under scrutiny in the present chapter provides a definite or convincing answer to this intriguing phenomenon. Let's go back to the starting point of the current chapter: It is clear from example (1) that Mag is a learner who has created her own idiosyncratic verbal system, which includes endings reminiscent of target verb forms. What is not known is the extent to which she is aware of the grammar rules that enable the French native speakers to differentiate, for example, an infinitive and a past participle. She started learning French at school, age 15. She spent 6 months as an au pair in Paris when she was 18. She had been at a French university for 3 months when her interlanguage was recorded. She might have been taught the relevant rules in an instructed setting but not been able to access them when producing oral discourse, thus displaying a lack of proceduralization of the rules. Or, she might not have understood the rules in a correct fashion and therefore have opted for a creative solution that is compatible with a variety of interpretations on the part of her interlocutor. Or, she might not have been taught those particular

sets of rules and rely on frequency effects from the input. It would have been interesting to ask the participant herself what she thought these forms stood for, consequently involving her in data construction and interpretation, as suggested in Revesz (2012). The use of a think aloud protocol (Leow & Morgan-Short 2004) could help gain access to speakers' representations and what they have in mind when using [e] forms, yet it could only take place as a retrospective task, by asking them to transcribe their own production or by having the participant listen to their production and ask them to comment on what they think they meant by the use of such forms. Notwithstanding, we have no way of making sure that the participant's representation is stable and that they know how to explain it to the researcher (Norris & Ortega 2012; see also Gass & Mackey 2000 for learner introspection and retrospection techniques).

To fulfill SLA's objective of describing and understanding the dynamics of interlanguage development, we need adequate transcription procedures to propose a valid interpretative framework for data analysis. In that regard, the spoken-centric approach seems a good fit for the purpose. Finally, we need to thoroughly document data transcription and data coding procedures. I believe that bearing those issues in mind when designing experimental settings is crucial to provide meaningful research results and contribute to a sound description of learner language development.

## Acknowledgements

I wish to thank the series editors as well as the anonymous reviewer and my fellow co-editors for their detailed and insightful remarks on earlier versions of this paper. They helped me considerably improve the shape and content of this chapter.

## References

- Andersen, Roger W. & Yasuhiro Shirai. 1994. Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition* 16. 133–156.
- Bartning, Inge & Suzanne Schlyter. 2004. Itinéraires acquisitionnels et stades de développement en français L2. *French Language Studies* 14. 281–299.
- Bassetti, Benedetta, Paola Escudero & Rachel Hayes-Harb. 2015. Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics* 36. 1–6.

- Brissaud, Catherine, Jean-Pierre Chevrot & Pascale Lefrançois. 2006. Les formes verbales homophones en /E/ entre 8 et 15 ans: Contraintes et conflits dans la construction des savoirs sur une difficulté orthographique. *Langue Française* 151. 74–93.
- Brissaud, Catherine & Jean-Michel Sandon. 1999. L'acquisition des formes verbales en /E/ à l'école élémentaire et au collège: Entre phonographie et morphographie. *Langue Française* 124. 40–57.
- Cook, Vivian J. (ed.). 2002. *Portraits of the L2 user*. Clevedon: Multilingual Matters.
- David, Jacques, Odile Guyon & Catherine Brissaud. 2006. Apprendre à orthographier les verbes: Le cas de l'homophonie des finales en /E/. *Langue Française* 151. 109–126.
- Detey, Sylvain & Jean-Luc Nespoulous. 2008. Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua* 118. 66–81.
- Dewaele, Jean-Marc. 2003. Compte rendu – hommage: l'œuvre de L. Selinker. 49. 153–159.
- Eubank, Lynn. 1993. On the transfer of parametric values in L2 development. *Language Acquisition* 3(3). 183–208.
- Fayol, Michel & Sébastien Pacton. 2006. L'accord du participe passé: Entre compétition de procédures et récupération en mémoire. *Langue Française* 151. 59–73.
- Flavell, John H. 1977. *Cognitive development*. Englewood Cliffs, NJ: Prentice Hall.
- Gass, Susan M. & Alison Mackey. 2000. *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum.
- Granget, Cyrille. 2015. Pourquoi le présent n'est-il pas si simple d'un point de vue psycholinguistique? *Language, Interaction and Acquisition* 6(1). 107–148.
- Gretsch, Petra & Clive Perdue. 2007. In F. Planck & I. Nikolaeva (eds.), *Finiteness: Theoretical and empirical foundations*, 432–484. Finiteness in first and second language acquisition. Oxford.
- Han, ZhaoHong & Elaine Tarone. 2014. Introduction. In E. Tarone & Z. Han (ed.), *Interlanguage: Forty years later*, 1–6. Amsterdam: John Benjamins Publishing Company.
- Hawkins, Roger. 2004. The contribution of the theory of universal grammar to our understanding of the acquisition of French as a second language. *French Language Studies* 14. 233–255.
- Hendriks, Henriëtte. 1998. Comment il monte le chat? en grim pant! *Acquisition et Interaction en Langue Étrangère* 11. 147–190. <https://journals.openedition.org/aile/1454>.

- Herschensohn, Julia. 2001. Missing inflection in second language French: Accidental infinitives and other verbal deficits. *Second Language Research* 17(3). 273–305.
- Hickmann, Maya. 1982. *The development of narrative skills: Pragmatic and metapragmatic aspects of discourse cohesion*. Chicago, IL: University of Chicago. (Doctoral dissertation).
- Hickmann, Maya, Henriëtte Hendriks & Françoise Roland. 1998. Référence spatiale dans les récits d'enfants français: Perspective inter-langue. *Langue Française* 118. 104–123.
- Hickmann, Maya, Henriëtte Hendriks, Annie-Claude Demagny, Helen Engemann, Tatiana Iakovleva, Yinglin Ji, Anna-Katharina Ochsenbauer & Efstathia Soroli. 2011. *La représentation de l'espace - études expérimentales et translinguistiques*. Manuel de codage.
- Hilton, Heather E. 2014. Oral fluency and spoken proficiency: Considerations for research and testing. In Pascale Leclercq, Amanda Edmonds & Heather E. Hilton (eds.), *Measuring L2 proficiency: Perspectives from SLA*, 27–53. Bristol: Multilingual Matters.
- Jaffré, Jean-Pierre. 2006. Pourquoi distinguer les homophones? *Langue Française* 151. 25–40.
- Jefferson, Gail. 1996. A case of transcriptional stereotyping. *Journal of Pragmatics* 26(2). 159–170.
- Keysar, Boaz, Shuhong Lin & Dale J. Barr. 2003. Limits on theory of mind use in adults. *Cognition* 89. 25–41.
- Klein, Wolfgang. 1994. *Time in language*. Routledge. London.
- Klein, Wolfgang. 2006. On finiteness. In Veerle Van Geenhoven (ed.), *Semantics in acquisition*, 245–272. Dordrecht: Springer.
- Klein, Wolfgang & Clive Perdue. 1989. The learner's problem of arranging words. In Elizabeth Bates & Brian MacWhinney (eds.), *The cross-linguistic study of sentence processing*, 292–327. Cambridge: Cambridge University Press.
- Kormos, Judit. 2006. *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Lardièrre, Donna. 1998. Dissociating syntax from morphology in a divergent L2 end-stage grammar. *Second Language Research* 14(4). 359–375.
- Lardièrre, Donna. 2012. Linguistic approaches to second language morphosyntax. In Susan M. Gass & Alison Mackey (eds.), *The Routledge handbook of second language acquisition*, 106–126. Oxon: Routledge.

- Leclercq, Pascale. 2011. *How do learners of French and English include space and time reference in narratives?* Paper presented at the EUROSLA 21 conference, Stockholm, 8-10 September 2011.
- Lenart, Ewa & Clive Perdue. 2004. L'approche fonctionnaliste: Structure interne et mise en œuvre du syntagme nominal. *Acquisition et Interaction en Langue Etrangère* 21. 85–121.
- Leow, Ronald P. & Kara Morgan-Short. 2004. To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition* 26(1). 35–57.
- Mackey, Alison & Susan M. Gass. 2012. *Research methods in second language acquisition: A practical guide*. Chichester: Blackwell Publishing.
- MacWhinney, Brian. 2000. *The CHILDES project: Tools for analysing talk*. Mahwah, NJ: Lawrence Erlbaum.
- Marsden, Emma, Rosamond Mitchell, Sarah Rule & Florence Myles. 2002. *Oral French interlanguage corpora: Tools for data management and analysis*. Occasional Paper No. 58.
- Mondada, Lorenza. 2000. Les effets théoriques des pratiques de transcription. *Linx* 42. 131–146.
- Mondada, Lorenza. 2002. Transcrire l'interaction. *Cahiers de Praxématique* 39. 45–75.
- Mondada, Lorenza. 2007. Commentary: Transcript variations and the indexicality of transcribing practices. *Discourse Studies* 9(6). 809–821.
- Myles, Florence, Janet Hooper & Rosamond Mitchell. 1998. Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning* 48(3). 323–364.
- Nickerson, Raymond S. 1999. How we know - and sometimes misjudge - what others know: Imputing one's own knowledge to others. *Psychological Bulletin* 125(6). 737–759.
- Nimz, Katharina & Ghada Khattab. 2019. On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research* 28. 387–413.
- Norris, John M. & Lourdes Ortega. 2012. Assessing learner knowledge. In Susan M. Gass & Alison Mackey (eds.), *The Routledge handbook of second language acquisition*, 573–589. Oxon: Routledge.
- Noyau, Colette, Et-Tayeb Houdaïfa, Marie-Thérèse Vasseur & Georges Daniel Véronique. 1995. The acquisition of temporality in French. In Rainer Dietrich, Wolfgang Klein & Colette Noyau (eds.), *The acquisition of temporality in a second language*, 145–210. Amsterdam: John Benjamins Publishing Company.



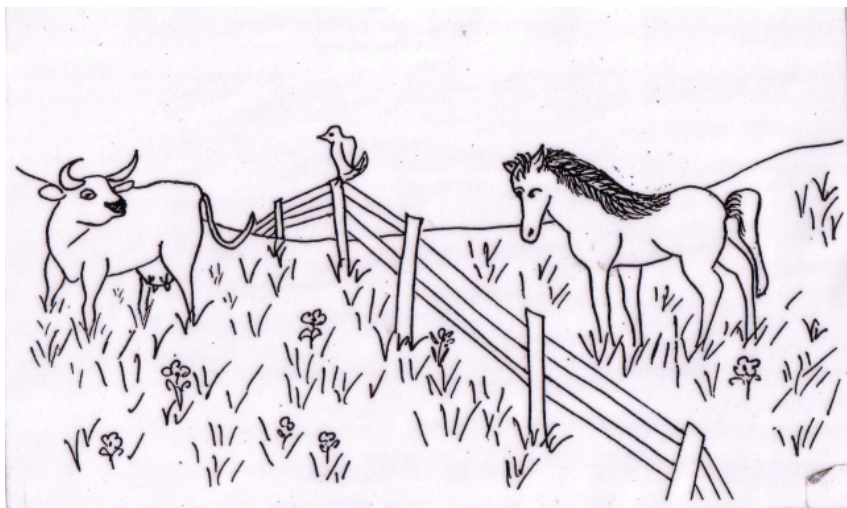
- Ochs, Elinor. 1979. Transcription as theory. *Developmental pragmatics* 10(1). 43–72.
- Ortega, Lourdes. 2014. Trying out theories on interlanguage: Description and explanation over 40 years of L2 negation research. In Zhao-Hong Han & Elaine Tarone (eds.), *Interlanguage: Forty years later*, 173–201. Amsterdam: Benjamins.
- Pallotti, Gabriele. 2017. Applying the interlanguage approach to language teaching. *International Review of Applied Linguistics* 55(4). 393–412.
- Perdue, Clive, Sandra Benazzo & Patrizia Giuliano. 2002. When finiteness gets marked: The relation between morphosyntactic development and use of scopal items in adult language acquisition. *Linguistics* 40(4). 849–890.
- Prévost, Philippe. 2007a. Knowledge of morphology and syntax in early adult L2 french: Evidence for the missing surface inflection hypothesis. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (eds.), *The role of formal features in second language acquisition*, 353–377. Mahwah, NJ: Lawrence Erlbaum.
- Prévost, Philippe. 2007b. L'influence du paradigme verbal de la langue maternelle sur la variabilité morphologique en acquisition du français L2 par des apprenants adultes. *French Language Studies* 17. 49–80.
- Prévost, Philippe & Lydia White. 2000. Missing surface inflection or impairment in second language? Evidence from tense and agreement. *Second Language Research* 16(2). 103–133.
- Revesz, Andrea. 2012. Coding second language data validly and reliably. In Alison Mackey & Susan M. Gass (eds.), *Research methods in Second Language Acquisition: A practical guide*, 203–221. Chichester: Blackwell Publishing.
- Robison, Richard E. 1995. The aspect hypothesis revisited: A cross-sectional study of tense and aspect marking in interlanguage. *Applied Linguistics* 16. 344–370.
- Rohde, Andrea. 1996. The aspect hypothesis and the emergence of tense distinction in naturalistic L2 acquisition. *Linguistics* 34. 1115–1137.
- Saturno, Jacopo. 2015. La trascrizione di varietà di apprendimento iniziali. In Marina Chini (ed.), *Il parlato in (italiano) L2: Aspetti pragmatici e prosodici. [Italian L2 spoken discourse: Pragmatic and prosodic aspects]*, 117–138. Milano: FrancoAngeli.
- Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the full transfer/full access model. *Second Language Research* 12(1). 40–72.
- Segalowitz, Norman. 2010. *Cognitive bases of second language fluency*. New York, NY: Routledge.
- Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics* 10. 209–231.

- Vainikka, Anne & Martha Young-Scholten. 1996. Gradual development of L2 phrase structure. *Second Language Research* 12(1). 7–39.
- Vainikka, Anne & Martha Young-Scholten. 1998a. Morphosyntactic triggers in adult SLA. In Maria-Luise Beck (ed.), *Morphology and its interfaces in second language knowledge*, 89–114. Amsterdam: John Benjamins.
- Vainikka, Anne & Martha Young-Scholten. 1998b. The initial state in the L2 acquisition of phrase structure. In Suzanne Flynn, Gita Martohardjono & Wayne O'Neil (eds.), *The generative study of second language acquisition*, 17–34. Mahwah, NJ: Laurence Erlbaum.
- Véronique, Daniel. 2004. The development of referential activities and clause-combining as aspects of the acquisition of discourse in French as L2. *French Language Studies* 14. 257–280.

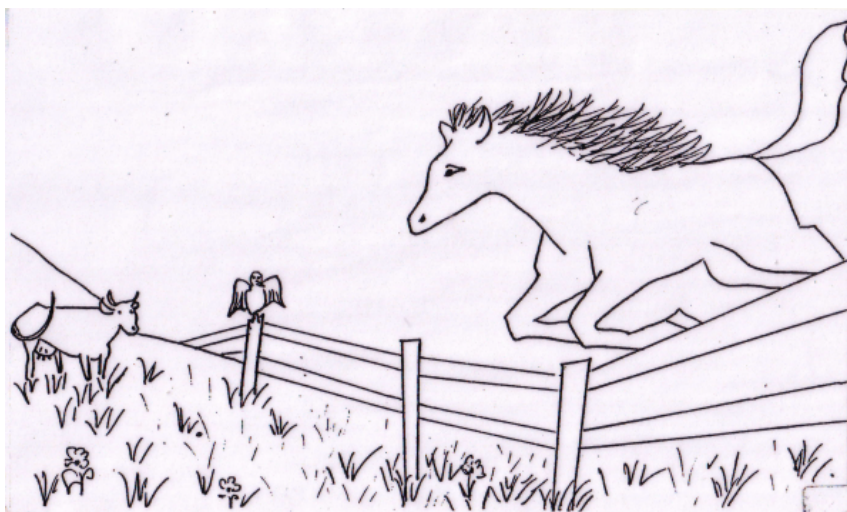
## Appendix: The horse story stimulus (Hickmann 1982)



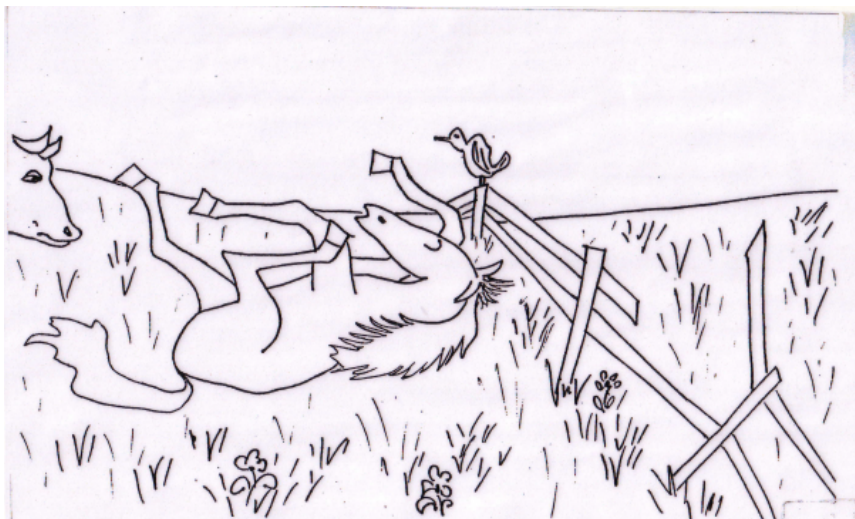
Picture 1



Picture 2



Picture 3



Picture 4



Picture 5