The role of aspect on anaphora resolution in English as a first and second language

ROBERTO B. SILEO (iD)

LUCA CILIBRASI (iD)

JULIA HEINE (iD)

IANTHI M. TSIMPLI (iD)

*Author affiliations can be found in the back matter of this article

ABSTRACT

This study investigates pronominal reference assignments across sentences that contain English verbs of transfer in monolingual English speakers and second-language (L2) learners having German as a first language and English as an L2. In a forced-choice task, participants were presented with sentences in perfective or imperfective aspect, like "Elizabeth took/was taking a meal to Mary" (adapted from Ferretti et al., 2009). They were then shown sentences that contained gender-matching pronouns, as in "She breathed in the smell of fresh basil", and they were finally asked to choose who performed the relevant actions: "Who breathed in the smell of fresh basil? Elizabeth or Mary?". We found that both groups preferred more often goal-oriented interpretations in the perfective condition, while in the imperfective condition only English monolingual speakers preferred more often source-oriented interpretations. The pattern observed in the perfective condition is consistent with previous studies and indicates that perfective aspect creates a strong bias towards end-states. For the imperfective condition, we argue that the different pattern observed in L2 learners may be due to some features of German, where an overall bias for end-states was previously observed. This indicates an effect of first-language strategies on L2 processing, consistent with previous research on different languages.

CORRESPONDING AUTHOR:

Luca Cilibrasi

University of Siena, IT;
Charles University, Prague, CZ

luca.cilibrasi@unisi.it

1. INTRODUCTION

Anaphoric pronoun resolution across ambiguous verb-of-transfer sentences is claimed to be affected by a range of linguistic features. While grammatical subjects have traditionally been seen as the most prominent candidates for antecedents (Broadbent, 1973; Crawley et al., 1990; Frederiksen, 1981; Hobbs, 1976; Rondal et al., 1984), this view is not consistently supported, and verb semantics (Garvey & Caramazza, 1974; Garvey et al., 1974–1975), thematic roles (Stevenson et al., 1994), and/or aspect (Ferretti et al., 2009; Kehler et al., 2008) have all been claimed to play a role in the interpretation of pronouns. In off-line story continuation studies in which participants were shown prompt sentences and asked to complete stories in their own time, Ferretti et al. (2009) and Kehler et al. (2008) revealed that aspect plays an all-important part in the interpretation of sentences with English verbs of transfer. These studies show that perfective aspect tends to trigger goal-oriented pronoun resolutions, while imperfective aspect tends to trigger source-oriented interpretations.[1] Aspect is then claimed to constrain interpreters' representation of events and, as a result, to also constrain anaphoric pronoun resolution.

2. BACKGROUND

2.1. SUBJECTHOOD AND OTHER LINGUISTIC FEATURES IN ENGLISH

It has traditionally been claimed that subjects are the most prominent antecedents in anaphoric pronoun resolution (Broadbent, 1973; Crawley et al., 1990; Frederiksen, 1981; Hobbs, 1976; Rondal et al., 1984). For example, when participants made reference assignment selections in a three-sentence self-paced reading task, Crawley et al. (1990) found that subject-oriented interpretations in ambiguous contexts like (1) were obtained more often than object-oriented resolutions. In the illustration below, Brenda was the preferred antecedent when participants were requested to answer "Who did Bill watch?".

(1)   Brenda and Harriet were starring in the local musical. Bill was in it too and none of them were very sure of their lines or the dance steps. Brenda copied Harriet and Bill watched her.
        (Crawley et al., 1990, p. 250)

The all-pervasive role of subjects has also been stressed by the so-called Centering Theory (e.g., Brennan et al., 1987; Grosz et al., 1995; Walker et al., 1998). Within that framework, as Walker et al. (1998) pointed out, all referents in an utterance are ranked according to salience and, in English, subjects are more salient than objects as antecedents in subsequent discourse and, in turn, objects are more salient than other grammatical roles (pp. 3–7). To illustrate, in (2) and (3), Susan appears as the preferred antecedent for the pronouns in both (d) and (e) examples.

(2a)   Susan is a fine friend.
(2b)   She gives people the most wonderful presents.
(2c)   She just gave Betsy a wonderful bottle of wine.
(2d)   She told her it was quite rare. (Susan told Betsy).
(2e)   She knows a lot about wine. (Susan knows…).

(3a)   Susan is a fine friend.
(3b)   She gives people the most wonderful presents.
(3c)   She just gave Betsy a wonderful bottle of wine.
(3d)   She told her it was quite rare. (Susan told Betsy).
(3e)   Wine collecting gives her expertise that's fun to share. (Susan's expertise).
        (Grosz et al., 1995, pp. 212–213)

While Centering Theory's explanation adequately accounts for the privileged role of the subject in each of (2) and (3), verb semantics and thematic roles have all been evidenced to affect pronominal reference assignment as well. First, in a story continuation task, Garvey and Caramazza (1974) demonstrated that there is more to anaphoric reference assignment than

---

1   In this paper, we use the terms "source(s)" and "goal(s)" as generally defined in the literature (e.g., Andrews, 1985; Fillmore, 1968; Radford, 1988). We use the terms "perfective" and "imperfective" to refer, respectively, to completed and ongoing events (e.g., Montrul & Slabakova, 2003).

a purely syntactic subject bias. The authors showed, for instance, that verbs' implicit causality relations led to preferences for subject-oriented interpretations in (4) and for object-oriented interpretations in (5). In (4), the cause of a confession primarily lies with the subject, the confessor, while in (5) the cause of the punishment primarily lies with the object, the daughter.

(4)   The prisoner confessed to the guard because he…
(5)   The mother punished the daughter because she…
       (Garvey & Caramazza, 1974, p. 461)

For the purposes of our analysis it is pertinent to point out that Garvey and Caramazza's (1974) participants did not reveal any subject- or object-oriented preference in perfective contexts that displayed the verb "to give". In other words, there were no preferred responses for any of the interpretations depicted in (6b) and (6c).

(6a)   Why did John give Walt the book?
(6b)   Because he … didn't need it anymore.
(6c)   Because he … wanted to read it.
       (Garvey & Caramazza, 1974, p. 461)

Garvey and Caramazza (1974) hypothesized that "to give" was either neutral in terms of implicit causality or more easily affected by other linguistic features, but there is, in this respect, one observation to make. Answers to open-ended "Why?" questions like (6a) cannot but be expected to be open-ended. Crawley et al. (1990) suggested that when questions are asked, anaphoric pronoun resolution is likely triggered by interpreters' conscious inferential thinking. Additionally, the process may be taken to be guided by the imaginary worlds that readers and/or hearers create when performing the task.

Thematic roles affect anaphoric reference resolutions too. Two sentence continuation tasks by Stevenson et al. (1994) looked into the interpretation of source-goal and goal-source experimental items. In Experiment 1, participants encountered the pronouns in a separate sentence, as in (7) and (8), while, in Experiment 2, participants encountered the pronouns in the sentence that contained the antecedents, as in (9) and (10).[2] Examples (7) and (9) depict sources in the subject, while examples (8) and (10) reveal sources in a non-subject position.

(7)   John passed the comic to Bill. He…
(8)   John seized the comic from Bill. He…
(9)   John passed the comic to Bill and he…
(10)   John seized the comic from Bill and he…
       (adapted from Stevenson et al., 1994, p. 525)

Stevenson et al. (1994) found that goals were generally preferred to sources in both experiments and that this was the case in goal-source and in source-goal target contexts. In (7) and (8), for instance, the pronouns in the second sentence were preferred to refer back to the goals, in both cases irrespective of the goals' grammatical position. Following Moens and Steedman's (1998) proposal, end-states may be considered the most salient component in the mental representation of events. According to the authors, an event such as "reaching the top of Mt. Everest" can be taken to comprise the preparatory processes of climbing steps and/or resting and the consequence or endpoint of being at the top (Moens & Steedman, 1998, p. 18).[3] With regard to endpoints, it is pertinent to mention telicity, the property of events which have inherent, lexically-encoded terminal points (e.g., Montrul & Slabakova, 2003). For example, the event of "passing a comic" is telic: It semantically involves the complete transfer of a comic from one person to another. Conversely, as Depraetere (1995) explained, the event of "sunbathing" does not have a similar inherent endpoint. A person's sunbathing may end after a while, but the termination of such an event does not itself reside within the meaning of "sunbathing" (Depraetere, 1995, p. 2). Montrul and Slabakova (2003) observed,

---

2   In a third experiment, Stevenson et al. (1994) tested story continuations after "so" and "because". They evidenced that in source-goal and goal-source sentences like "John passed the comic to Bill because… / so…" and "John seized the comic from Bill because… / so…", goals were generally preferred to sources when experimental items ended in "so".

3   It is also worth pointing out that Stevenson et al.'s (1994) end-state bias account technically implies, as Kehler et al. (2008) put it, that preferences for goal-oriented interpretations in perfective contexts are an epiphenomenon.

in this respect, that both telic and atelic verbs could be used in perfective and imperfective contexts. For instance, while the perfective use of the telic verb of transfer in (7) shows that the event has reached an actual temporal endpoint, the same telic expression can also be used in imperfective or ongoing settings. This makes aspect a relevant variable to test in itself and, in the following section, we discuss two studies that form the basis of our empirical research in this paper. Both such studies focus on English verbs of transfer and the all-important role of aspect.

2.2. ASPECT AND ANAPHORA

Aspect, the linguistic category distinguishing complete from incomplete events (e.g., Montrul & Slabakova, 2003), has also been claimed to affect anaphoric pronoun resolution. Kehler et al. (2008) found that reference assignment across sentences with English verbs of transfer is sensitive to perfective and/or imperfective aspect. In a story continuation task, Kehler et al. (2008) investigated the prevalence of goal- and source-oriented interpretations in perfective and imperfective contexts. Each experimental item contained a full sentence depicting a source and a goal of the same gender, followed by the second sentence that participants had to complete, which was introduced by a pronoun. For example, participants were presented with items such as (11) or (12).

(11)   John was handing a book to Bob. He ...
(12)   John handed a book to Bob. He ...
         (Kehler et al., 2008, p. 24)

Participants were thus free to continue the story focusing either on the source or the goal of the previous sentence. Linguistic judges would later evaluate the responses, based on the selection of words used, and establish who the chosen referents were. Source completions represented a majority of participants' responses in both the perfective and imperfective conditions, which reveals an overall bias towards source-oriented reference assignments when using this methodology. However, this pattern was significantly stronger in the imperfective condition.[4]

In another story continuation task, Ferretti et al. (2009) found different yet compatible results. In the task, each experimental item consisted of one full sentence with sources and goals that differed in gender, as illustrated in (13) and (14), and participants had to write out an entire continuation sentence without any pronoun provided.[5]

(13)   John handed a book to Mary. ...
(14)   John was handing a book to Mary. ...
         (Ferretti et al., 2009, p. 193)

Unlike Kehler et al. (2008), Ferretti et al. (2009) found an overall bias towards goals in both perfective and imperfective contexts. However, in this case, perfective aspect was found to strengthen that bias, and the difference between the number of goal-oriented interpretations in perfective and in imperfective contexts was statistically significant.[6] Notably, Ferretti et al.'s (2009) event-related potential (ERP) experiment revealed compatible results with their offline story continuation data. In this second experiment, participants read ongoing or completed transfer of possession sentences in which the following pronouns matched either goals or sources in gender, as in "Sue handed/was handing a timecard to Fred. She/he asked about the upcoming meeting", while the electroencephalogram was recorded. Ferretti et al. (2009) reported that the goal bias evidenced in their other experiment was manifested differently in the brain in perfective and in imperfective contexts.

As mentioned, Ferretti et al. (2009) and Kehler et al. (2008) revealed partly compatible findings. Despite different overall trends, both studies showed a statistically significant bias towards goal-oriented interpretations in perfective contexts, and a statistically significant bias towards source-oriented interpretations in imperfective contexts. In Ferretti et al. (2009), participants

---

4   See Figure 1.

5   Ferretti et al. (2009) included a gender distinction in their design, first, to look into participants' expectations as to whether sources or goals would be referred to next and, second, to look into what referring expressions (pronouns and/or others) participants used in their completions.

6   See Figure 1.

chose goals overall more often in both aspects, but this pattern was less pronounced in the imperfective condition. In Kehler et al. (2008), participants chose sources overall more often in both aspects, but this pattern was less pronounced in the perfective contexts. To ease visualisation of the partly compatible and partly contrasting findings in these studies, we show Ferretti et al.'s (2009) and Kehler et al.'s (2008) results in Figure 1.
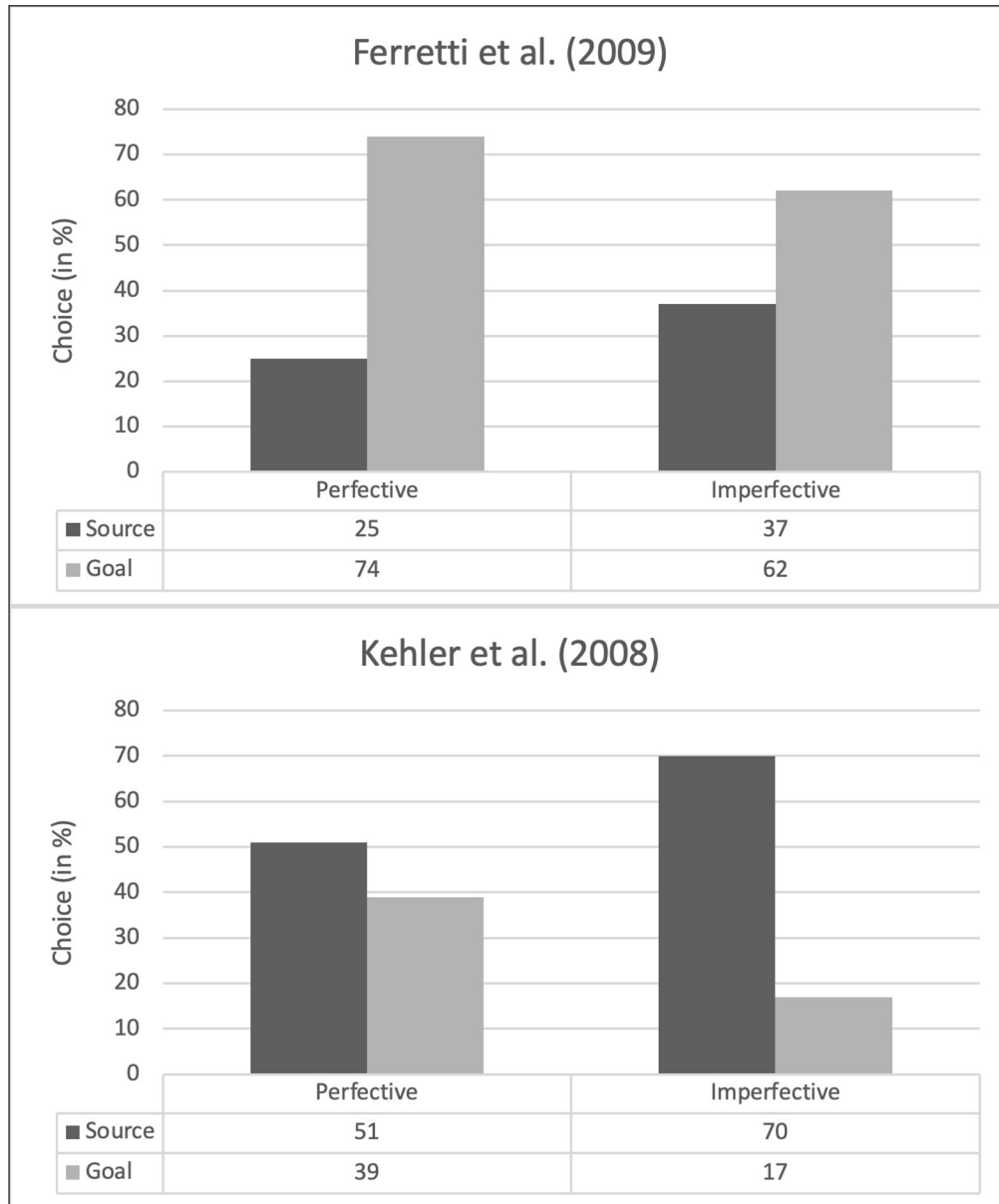
Ferretti et al. (2009)	Perfective	Imperfective
■ Source	25	37
■ Goal	74	62

Kehler et al. (2008)	Perfective	Imperfective
■ Source	51	70
■ Goal	39	17

Figure 1 Visual representation of the numerical data reported by Ferretti et al. (2009) and Kehler et al. (2008).

This complex pattern of results leaves room for further research. While aspect appears to affect results in similar ways in the two studies, there also appears to be an inconsistency in participants' baseline preferences. It is possible that the different methodologies employed by Kehler et al. (2008) and Ferretti et al. (2009) are the origin of these contrastive patterns. In Ferretti et al. (2009), participants were given complete freedom on how to continue their sentences. In such context, it is possible that an overall preference for goal referents may be a reflection of recency effects. Recency effects are a classic finding in cognitive psychology and are usually described as a property of human working memory (Dosher, 2006): Information that is presented last tends to be more active and more easily recalled. When subjects were given complete freedom on what to recall, they may have chosen to resort to goals because these were more active in their working memory. In Kehler et al. (2008), participants were given the first pronoun, and thus the issue was deciding how to use this pronoun. The option was to take this either as a reference to the source or as a reference to the goal and then it was necessary to complete the sentence with relevant information. This type of task requires a more complex type of processing, so participants' interpretations are less prone to be affected by simple recency effects. An exploration of semantic representations, and a more conscious choice, are necessary to complete the sentence.

It is worth pointing out that Ferretti et al. (2009) and Kehler et al. (2008) required participants to freely complete story continuation tasks in their own time, a procedure which, as we further elaborate in section 6 could be taken to trigger inferential responses. One issue that also arises in both Ferretti et al. (2009) and Kehler et al. (2008), which suggests that subjects were overthinking, is that a substantial number of responses provided by participants did not clearly fit with either source or goal interpretations, and this led to a considerable amount of missing data. Ferretti et al. (2009) revealed, for example, that approximately 21% of 40 participants' responses had to be disregarded as they contained unclear or nonsensical continuations, fragmented answers, misinterpreted genders and/or no references to either goals or sources altogether (p. 194). Likewise, in Kehler et al. (2008), the authors stressed that they were not able to disambiguate all answers (p. 25). They showed, on the one hand, that 10% of responses were ambiguous in the perfective aspect, where the distance between goal versus source was smaller (and as such a 10% of ambiguous data is an issue), and, on the other hand, that 17% of responses were ambiguous in the imperfective aspect, where the bias towards source was instead strong (and the ambiguous responses were less of a problem, even though the percentage was rather high). As it becomes apparent in section 4.2, in our study we aim at overcoming this issue. In our task, participants were not allowed to go forward without providing their answers, so choosing either goal or source interpretations was obligatory. In addition, having to select one of two possible answers in our forced-choice task may have led to more intuitive and less inferential responses, in comparison to previous studies.[7]

2.3. ANAPHORIC REFERENCE ASSIGNMENT IN SECOND-LANGUAGE (L2) LEARNERS

Previous work has shown that bilinguals and L2 learners may be influenced in their anaphora resolution by features of their first language (L1). Typically, these effects have often been shown to be related to whether the two languages differ in the presence of null subjects. When using their L1, speakers of languages that allow null subject pronouns treat null and overt subject pronouns differently: While null pronouns tend to be assigned to a salient referent (usually the first available explicit subject), overt pronouns tend to be assigned to a less-salient referent more often. As Cunnings et al. (2017) explain, in null subject languages, overt pronouns are used to elicit a topic-shift, meaning that they move the attention to a less obvious referent than the default one assigned with the null pronoun. Within the realm of L2 English research, interesting results were recently obtained by Contemori et al. (2019) testing Mexican Spanish learners of English, where only Spanish licenses null subject pronouns (note that similar results were previously obtained by Roberts et al. (2008) with different language pairs). In this study, when presented with English subject pronouns, Mexican Spanish learners of English chose non-local referents more often than English native-speakers, mirroring strategies adopted in anaphora resolution in Spanish when presented with an overt pronoun. Jegerski et al. (2011) tested the opposite group: Here, intermediate and advanced learners of Spanish with English as an L1 failed to capture the distinction in the use of null and overt pronouns, suggesting a lingering L1 effect on L2 anaphora resolution.

Despite these important differences, some studies have also found that L2 learners may reach reliable native-like performance in anaphora resolution. Cunnings et al. (2017) tested Greek learners of English and a control group of native speakers in anaphora resolution contexts disambiguated by gender. In this experiment, both groups performed very accurately and both showed to quickly be able to use gender to resolve anaphoric assignment. These results were obtained despite the crucial difference between English and Greek regarding the use of null subjects and regarding topic-shift, which is typical of overt subjects in Greek but not in English.

Previous research has also demonstrated successful L2 learning of peculiar properties. The findings presented in this paragraph are connected to a relevant phenomenon described in the literature as the overt-pronoun constraint (OPC; Montalbetti, 1984). The OPC states that "in null subject languages, overt pronoun subjects of embedded clauses cannot be bound by wh- or quantifier antecedents" (Gelormini-Lezama et al., 2016, p. 979). For instance, in the Spanish sentence *Muchos plomeros creen que ellos compraron un pulpo* 'Many plumbers believe that they bought an octopus', the pronoun "they" cannot refer to "plumbers". On the contrary,

---

[7]   See section 4.3.

when "they" is substituted with the null pronoun, as in *Muchos plomeros creen que compraron un pulpo*, the pronoun may refer to "plumbers" (Montalbetti, 1984, p. 78). Several studies on L2 acquisition suggest that learners may readily access this constraint, even when their L1 belongs to a different category than the L2 (Buckledee, 2008; Kanno, 1997; Pérez-Leroux & Glass, 1999).

Interestingly, Contemori and Dussias (2016) showed that when L2 learners of English are prompted to produce sentences with anaphoric expressions, their outputs may differ from those of native English speakers, although these differences may not be predicted by their L1. In this study, Spanish learners of English and English native speakers were asked to retell a story they heard or create a story based on pictures they were presented with. The analysis of anaphoric expressions showed that while native speakers tended to use full noun phrases to reduce ambiguity, L2 learners tended to use pronouns. This choice is not predicted by L1-L2 interference, but it rather suggests that L2 speakers opt for the choice that is less costly for them. The use of a pronoun might be a default option that L2 learners employ to avoid the cost of finding the exact words to identify the referent they need.

Another explanation that refers to the cost of processing in the L2, and to the strategies deployed to avoid these costs, may be related to the so-called Shallow Structure Hypothesis (SSH). This hypothesis suggests a tendency of L2 processing to rely more strongly on nongrammatical information than L1 processing (Clahsen & Felser, 2017). Crucially, and opposed to some researchers' interpretation of the initial version of the SSH (Clahsen & Felser, 2006), Clahsen and Felser (2017) emphasized that this stronger reliance on pragmatic and semantic information in the L2 represents a gradual difference. An increasing exposure to the L2 will lead to a more automatic access to grammatical information in the L2. Nevertheless, Clahsen and Felser (2017) argued that this automatic processing of grammatical information remains less likely in L2 processing even in speakers with high proficiency levels. Based on these ideas, the authors suggested that, whereas the SSH does not rule out an interlinguistic influence of the L1, a general tendency to rely more strongly on nongrammatical information while underusing grammatical information should predict a less significant role of L1 transfer than had previously been suggested. In consideration of such an account, a stronger reliance on superficial information in the L2 could be assumed to lead to more pronounced domain-general effects of information processing such as recency effects. In our task, this would entail a general preference for goal referents (them being closer), irrespective of L1 features.

In summary, while some studies indicate an effect of L1 over L2 in anaphora resolution strategies, other studies fail to find these effects. Most of these studies addressed pairs of languages that differ in the presence of null subjects, because this property leads to a very different system of anaphoric assignment. Our study, instead, deals with two languages that do not differ in this sense, as both require explicit subjects. This choice offers ground to investigate whether other properties that do differentiate these two languages may have an effect on anaphoric resolution strategies.

In the current study, as we discuss in the sections that follow, we tested English L2 learners with German as L1. German is an interesting L1 to use for testing in this task because previous work has shown an endpoint bias during sentence processing for German learners of English where monolingual speakers do not display one. A crucial reference is a recent study by Roberts and Liszka (2021). In this study, German and French learners of English were presented with a self-paced reading task of garden path sentences in English. Sentences, as in (15), were built as to initially lead to a late-closure and eventually require an early-closure.

(15)   While John hunted the frightened rabbit escaped.
       (Roberts & Liszka, 2021, p. 1)

Participants were initially led to interpret "the frightened rabbit" as a complement of the verb "hunted", but they needed to eventually recalculate and interpret it as the subject of "escaped". The word "escaped" is the disambiguating point and, in self-paced reading tasks, it is shown to take a long time to be processed. In the group of English native speakers, there was evidence that aspect played a role, and the effect was significantly milder when the verb "to hunt" was presented in imperfective aspect, as in "While John was hunting the frightened rabbit escaped". In this study, Roberts and Liszka (2021) showed that while French learners of English behaved

like English natives, German learners did not. German learners were not heavily affected by aspect and displayed a bias for the late closure in both the perfective and the imperfective conditions. Roberts and Liszka (2021) suggested that this indicates a preference for endpoint interpretation of incoming material when Germans are processing their L2 due to an endpoint bias German leaners usually display in their L1. Importantly, because a comparable tendency was not observed for French participants, a general endpoint bias due to shallower processing strategies in the L2 can be ruled out as a sole explanation.

Roberts and Liszka (2021) explained this result referring to the grammaticalization of aspect. There appears to be a relationship with whether a language has grammaticalized aspect or not and the prevalence of endpoint preferences: Speakers of languages that do not have grammaticalized aspect, such as German, have a tendency for endpoint interpretation of upcoming material when using their first language (Carroll et al., 2008). In their study, Roberts and Liszka (2021) showed that these effects were extended to the L2, where German learners behaved differently than monolingual English speakers when presented with English garden path sentences. These findings may be relevant for our study, as they indicate an object bias when interpreting new material: The noun phrase "the frightened rabbit" in the example above tends to be interpreted as a complement of the verb "was hunting" more often by German learners than by French learners and English natives.

It should, however, be noted that an alternative explanation, which is not discussed in Roberts and Liszka (2021), is related to orthographic rules. In German, orthographic rules strictly require a comma after the verb "to hunt" if "the frightened rabbit" is to be interpreted as the subject of a new clause. In the absence of this comma, German speakers may be strongly biased to interpret "the frightened rabbit" as the object of the verb "to hunt", independently of aspect. Indeed, studies have shown that commas can directly contribute to the parsing and the disambiguation of sentences in on-line processing (e.g., Steinhauer & Friederici, 2001; Kerkhofs et al., 2008). Furthermore, it has been demonstrated that some languages, particularly German, "have stronger, or maybe more consistent, punctuation styles that interfere in the production of L2" (Markov et al., 2018, p. 3463), even when a high proficiency in the L2 has been attained. The same influence of punctuation was not observed for French, which could potentially explain why no comparable endpoint preference was observed for French participants in Roberts and Liszka's (2021) study either. The present study has the potential to discriminate between the influence of a general L1 endpoint bias and the role of punctuation in experimental stimuli, as the present target sentences did not involve such punctuation-induced conflicts.

3. RESEARCH QUESTIONS AND PREDICTIONS

Story continuation tasks with English native speakers provide us with a picture in which (i) event consequences/end-states appear to be salient components in the mental representation of perfected events (Stevenson et al., 1994), (ii) goals are preferred in both perfective and perfective verb-of-transfer contexts when no pronoun is provided, though the bias is less prominent in imperfective aspect (Ferretti et al., 2009), and (iii) sources are preferred in both perfective and imperfective verb-of-transfer contexts when a pronoun is given, though the bias is less prominent in perfective aspect (Kehler et al., 2008). Based on these studies, our hypothesis was that aspect would have a significant effect on anaphoric reference assignment. More specifically, we hypothesized that perfective aspect would predominantly lead to goal-oriented interpretations, while imperfective aspect would predominantly encourage source-oriented reference assignments. In order to investigate on-the-go interpretations in source- and goal-oriented pronominal resolutions in perfective and imperfective contexts, we departed from story continuation studies and used a forced-choice task. In addition to testing English monolingual speakers, we investigated whether identical patterns could be observed in German L2 learners of English. The current study therefore aims at answering the following questions:

a. When facing anaphora resolution in a forced-choice task, do English monolinguals display an overall preference for goals (as in Ferretti et al., 2009) or for sources (as in Kehler et al., 2008)?

b. Do German learners of English differ from English monolinguals in their anaphora resolution strategies in either perfective or imperfective contexts?

Our predictions were that both groups would display a preference for object/goal interpretations in the perfective condition and subject/source preferences in the imperfective condition, even though we expect the effect to be significantly milder for German L2 learners in the imperfective condition due to their baseline object preference.

4. METHOD

4.1. PARTICIPANTS

For the L1 group, 42 undergraduate and graduate students from the University of Cambridge participated in the task (23 women, 19 men). All participants were monolingual speakers of British English and over 18 years of age ($M = 20.05$, $SD = 2.15$). No participant was involved in linguistics and/or modern languages studies. Every participant signed a customary participation consent form and received £6 as compensation.

L2 participants were recruited online through the platform *Prolific* indicating the following selection criteria: age of onset ranging from 4 to 13 years, mostly grown up in Germany, currently living in Germany, English proficiency of C1–C2,[8] German as a native language, aged 18–40, no relevant impairments, and regularly speaking English. Forty-five L2 learners were finally included in the study (25 women, 19 men, 1 other). All participants were above 18 years of age ($M = 26.77$, $SD = 5.2$) and had started learning English during childhood ($M = 8.54$ years; $SD = 2.50$). Every participant received £2.50 for their participation through the platform *Prolific*.

4.2. MATERIALS

We showed participants perfective and imperfective sentences with sources and goals of the same gender, each followed by a second sentence containing a pronoun. We then asked, following Crawley et al.'s (1990) method, "Who did what?" questions. Like Kehler et al. (2008), we introduced all sources in subjects. In other words, we combined elements from precedent studies and adapted the method in order to address our aims.

The experimental items consisted of 28 targets and 28 distractors. We adapted the targets from Ferretti et al.'s (2009) ERP task.[9] Ferretti et al.'s (2009) stimuli contained sentences with perfective and imperfective verbs of transfer followed by sentences that could trigger source- or goal-oriented interpretations, as in (16a).

(16a)   Elizabeth was taking/took a meal to Frank. He/She breathed in the smell of fresh basil. (Ferretti et al., 2009, Appendix A)

Because we looked into cross-sentential reference assignment irrespective of gender, we replaced proper names in each goal and kept pronouns consistent, in terms of gender, in each second sentence. While the second sentences in Ferretti et al. (2009) began with a pronoun that unambiguously referred either to the source or the goal in the first sentence, the pronouns in our targets could trigger ambiguous source/goal interpretation, as in (16b).

(16b)   Elizabeth was taking/took a meal to Mary. She breathed in the smell of fresh basil. (adapted from Ferretti et al., 2009, Appendix A)

It is pertinent to point out that Ferretti et al.'s (2009) items in the perfective condition entailed that the relevant events had finished, while the imperfective items in the study might have triggered either complete or incomplete event interpretations. In our study, a similar observation can be made, this time across the two relevant sentences for each item. The perfective form in the first sentence in (16b) entails completion of the relevant event and having received the meal, Mary was in a position to breathe in the smell of fresh basil. The imperfective form in the first sentence in (16b), on the contrary, does not entail that the event had been finished and respondents could have well imagined either that the meal had or had not finally reached Mary.

---

8   An explanation of the proficiency levels according to the Common European Framework of Reference for Languages was provided to all participants.

9   We did not mention this experiment in the introduction because the use of unambiguous sentences makes it irrelevant for our discussion. However, with small changes, Ferretti et al.'s (2009) trials revealed to be optimal to investigate our research questions. We are grateful to Todd R. Ferretti for allowing us, on behalf of his team, to adapt and use some of the stimuli from their study.

After all, the smell of fresh basil could also have been breathed in from a distance. Despite this uncertainty in imperfective contexts, the key point for the purposes of our task was that the pronouns in all second sentences in our items allowed for ambiguous interpretations in both perfective and imperfective contexts.

In order to test each target in perfective and in imperfective aspect, and in order to avoid the presentation of the same targets, albeit in different aspects, to the same participants, we created two lists of items. Each list was made of 14 perfective and 14 imperfective sentences. If a sentence was in perfective aspect in list A, it would be in imperfective aspect in list B (and vice-versa). Each participant was then presented with only one of the two randomized lists. All items are available in Appendix A. The 28 targets contained a balanced number of female and male proper names.

The 28 distractors also consisted of two sentences. The first sentence in each distractor contained sources and goals, as was the case in the targets, but no verb of physical transfer. The second sentence depicted events which were clearly related to either the source or the goal. As Hirst and Brill (1980) suggested, strong contextual constraints can ease pronominal reference assignments. For example, the pronoun in (17) clearly refers to John, while the pronoun in (18) could trigger ambiguous interpretations.

(17)   John stood watching while Henry fell down some stairs. He ran for a doctor.
(18)   John stood watching while Henry fell down some stairs. He thought of the future.
        (Hirst & Brill, 1980, p. 168)

In order to distract participants further, 14 distractors displayed perfective aspect, as in (19) and (20), while the remaining 14 were imperfective, as in (21) and (22).

(19)   Susan diverted the query to Martha. She expected her friend to give the right answer.
(20)   Vivian promised a promotion to Charlotte. She had been working very hard.
(21)   Oliver was making an early call to Benjamin. He enjoyed calling friends in the morning.
(22)   Esther was paying full attention to Valerie. She was saying interesting things.

The same distractors were used for subgroup A and subgroup B and, as was the case with the targets, distractors contained a balanced number of female and male proper names. Appendix A displays all targets and distractors in the task.

4.3. PROCEDURE

The task was designed and performed on *Psychopy* (v3.1.5). Monolingual speakers completed the task on a 13-inch Apple MacAir computer, while the L2 learners completed the task online on their own laptop or desktop computer through the platform *Psytoolkit* (Stoet, 2010).[10] They were informed that two sentences would appear on the computer screen, one at a time, and that "Who did what?" questions would have to be answered after the second sentence was shown. Participants in the monolingual group were instructed to wait for each question and to press either ○, which corresponded to sources, or ●, which corresponded to goals. There were ○ and ● stickers on letters "q" and "p" on the computer keyboard. The L2 learners, instead, completed the task using their keyboard, and thus the letter "p" and "q" without stickers. Participants encountered two practice items to start with and, once the practice was over, they pressed the space bar to begin. The three sentences corresponding to the first practice item in the task are shown in (23–25).

(23)   Diana was transferring $1000 to Laura.
(24)   She checked the bank's routing number.
(25)   Who checked the bank's routing number? Diana (○/p) or Laura (●/q)?
        (adapted from Ferretti et al., 2009, Appendix A)

All sentences were shown in the center of the computer screen. Once the first sentence appeared, it remained on display for four seconds. It was then replaced by the second sentence, which remained on the screen for the same length of time. Then the relevant question was

---

10   Interfaces and procedure were identical on *Psychopy* and *Psytoolkit*. The only difference was whether participants were tested in the presence of the experimenter (*Psychopy*) or whether they completed the task on their own (*Psytoolkit*). Inspection of the data and the actions we took to exclude outliers lead us to say that both data collections are equally valid.

shown until participants' ○/p or ●/q selections were made. No response time limit was set in either of the versions. The first sentence of the next experimental item was then presented and, after the practice, the sequence was repeated 56 times (for targets and distractors) without interruption.

5. RESULTS

5.1. MONOLINGUAL PARTICIPANTS

The task was completed by 42 monolingual participants in total, with 21 participants in two subgroups. Because participants were presented with each verb in one aspect only, they made 56 pronoun resolutions each, which yielded 1,176 responses per subgroup and a total of 2,352 responses. Half of the experimental items were distractors, so we collected 1,176 target responses and an equal amount of distracting responses. We first calculated the number of source and goal interpretations in perfective and imperfective aspect. This enabled us to compare the number of instances of each interpretation in both the perfective and the imperfective conditions. Responses that took more than three standard deviations from the mean to be provided were considered outliers and were then removed. Nineteen responses were removed in total, amounting to 1% of the dataset.[11] Descriptive statistics are presented in Table 1.

CHOICE	PERFECTIVE	IMPERFECTIVE
Source	43	56
Goal	57	44

Table 1 Choices across conditions (in %).

Data were then analyzed in R using logistic linear mixed models (lgmer), from the library lme4 (Bates et al., 2007). Because we were interested in understanding the role of aspect in predicting a goal/source bias, aspect was used as a predictor, and choice was used as the outcome variable. Both aspect and choice were dummy coded (zero versus one) and both were re-levelled using the sliding contrast function (from the MASS library). Following a well-established procedure (Baayen et al., 2008), the random structure was identified by comparing four random structures (the full list is available in Appendix B), and the one with the smallest Akaike Information Criterion was chosen. The final model was thus:

Choice ~ aspect + (1|part) + (aspect|item)

Results are presented in Table 2. The outcome of this model indicates that aspect was a highly significant predictor of choice.

FIXED EFFECTS	ESTIMATE	STANDARD ERROR	Z VALUE	p
Intercept	–0.08	0.29	–0.31	0.75
Aspect	–0.75	0.19	–3.93	<.001 ***
Random effects	**Estimate**			
σ^2	2.34			
$\tau_{00\,part}$	0.13			
$\tau_{00\,item}$	2.12			
τ_{11}	0.41			
ρ_{01}	0.14			

Table 2 Summary from the logistic regression model.

To understand the role of this predictor (i.e., the direction of the biases related to aspect), post-hoc analyses were performed. Post-hocs (chi-squares) showed that monolingual English speakers displayed a bias for goal interpretation in perfective sentences (χ^2 (1) = 10.8, p = .001),

---

11   Response times were used to exclude outliers, as both platforms provided them "for free". We do not report them, however, as our main objective was to focus on anaphora resolution choices.

while they displayed a bias for source interpretation in imperfective sentences (χ^2 (1) = 7.4, p = .006). These patterns are visualized in the bar-chart in Figure 2.
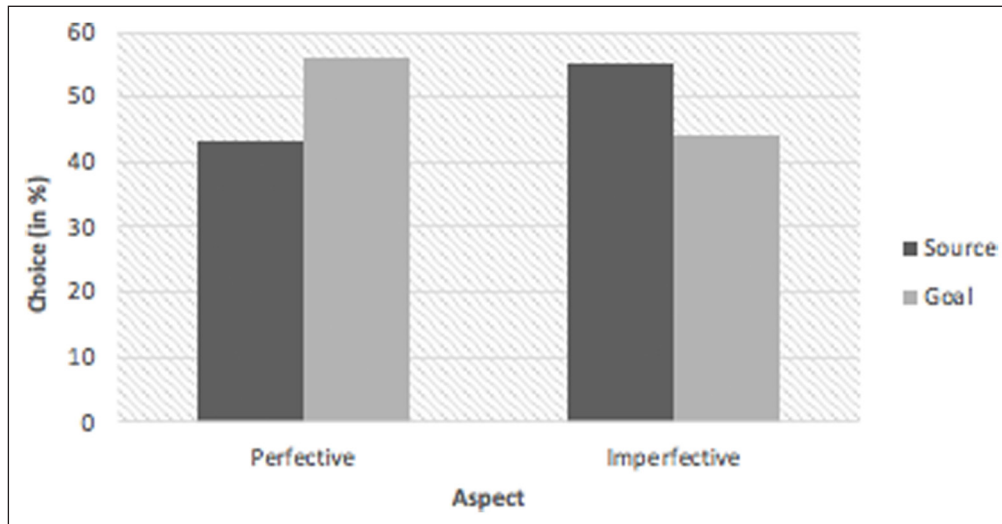
Figure 2 Source and goal interpretations across aspects in the monolingual English speakers.

5.2. L2 LEARNERS

The task was completed by 45 participants in total, with 22 participants in one subgroup and 23 in the other, making 56 pronoun resolutions each, which yielded a total of 2,520 responses. Half of the experimental items were distractors, so we collected 1,260 target responses and an equal amount of distracting responses. We first calculated the number of source and goal interpretations in perfective and imperfective aspect. This enabled us to compare the number of instances of each interpretation in both the perfective and the imperfective conditions. Responses that took more than three standard deviations from the mean to be provided were considered outliers and were then removed. Thirty-two responses were removed in total, amounting to 2% of the dataset. Descriptive statistics are presented in Table 3.

CHOICE	PERFECTIVE	IMPERFECTIVE
Source	38	47
Goal	62	53

Table 3 Choices across conditions (in %) in the English as an L2 group.

Data were then analyzed in R using logistic linear mixed models (lgmer) and following the same procedure adopted for the monolingual speakers. The best model revealed to be M1: Choice ~ aspect + (1|part) + (1|item).[12] The results indicate that, also in this case, aspect was a significant predictor of choice. Results are presented in Table 4.

FIXED EFFECTS	ESTIMATE	STANDARD ERROR	Z VALUE	p
Intercept	0.45	0.29	1.55	0.12
Aspect	–0.46	0.13	–3.45	<.001 ***
Random effects	**Estimate**			
σ^2	1.95			
τ_{00part}	0.34			
τ_{00item}	1.61			

Table 4 Summary from the logistic regression model.

To understand the role of this predictor (i.e., the direction of the biases related to aspect), post-hoc analyses were performed. Post-hocs (chi-squares) showed that L2 learners displayed a bias for goal interpretation in perfective sentences (χ^2 (1) = 35.7, p = .001), while they did not display significant biases in imperfective sentences (χ^2 (1) = 2.29, p = .13). These patterns are visualized in the bar-chart in Figure 3.
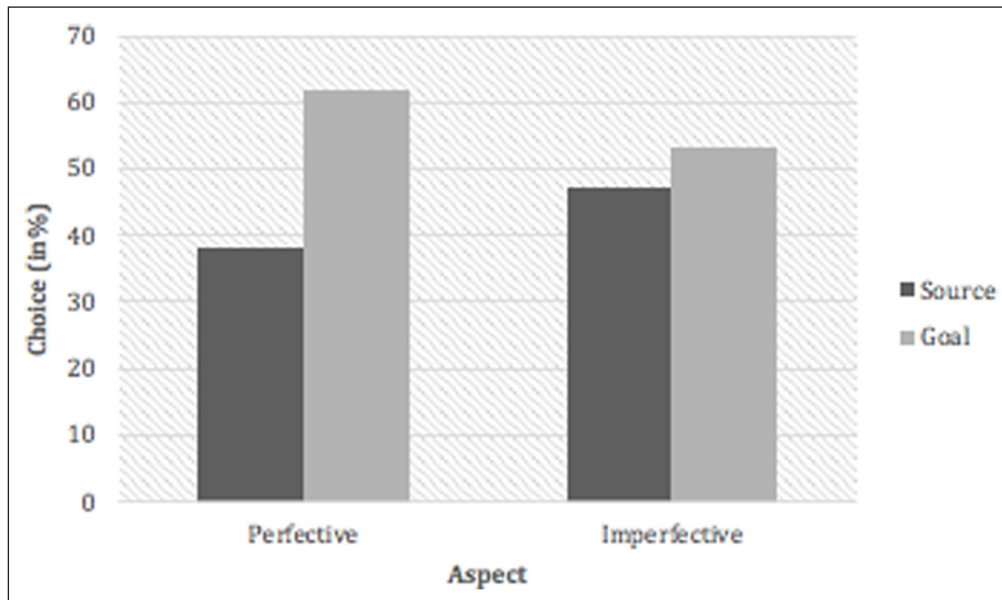
---

12   See Appendix B.

Figure 3 Results. Source and goal interpretations across aspects in the L2 learners of English.

6. DISCUSSION

Using a purposely developed forced-choice task, this study found that goals were significantly preferred in perfective verb-of-transfer contexts by both monolingual English speakers and L2 learners. Moreover, sources were significantly preferred in imperfective scenarios by English monolinguals, while German learners of English did not display any bias in the imperfective condition. We first discuss our results on monolinguals in comparison to previous work on the same population, and then we discuss the pattern observed in the L2 learners.

When it comes to the monolingual English speakers, our results are overall consistent with previous work, although some important differences are observed. While our results confirm Ferretti et al.'s (2009) and Kehler et al.'s (2008) findings on the role of aspect in that perfective aspect led to more goal-oriented interpretations and imperfective aspect triggered more source-oriented interpretations, contrary to these two previous studies, we did not find an overall preference for either goals or sources. We found, instead, opposite trends in each of the conditions we tested. Unlike Ferretti et al. (2009), who found an overall bias towards goals, and unlike Kehler et al. (2008), who found an overall bias towards sources, our results showed a trend for goals in the perfective condition and a trend for sources in imperfective contexts.

The different result we obtained may be related to methodological choices. Participants in Ferretti et al. (2009) and Kehler et al. (2008) were asked to imagine a story continuation for each target and to write out the first sentence that sprang into their minds. In Kehler et al.'s (2008) story continuation task, participants found a pronoun right after each target sentence. They were then asked to continue the sentence describing something relevant for that pronoun, and the words used were finally evaluated by judges to establish if the referent was either the source or the goal. With this method, as in all off-line tasks, there might have been room for substantial conscious inferential thinking. For instance, upon encountering "He" in each of (11) and (12) in section 2.2, participants may have engaged in a more conscious effort, in their own time, to assign a reference to the pronoun.

On the other hand, Ferretti et al.'s (2009) story continuation study, where participants were not given any word to start the new sentence, as illustrated in (13) and (14) in section 2.2, may have led to a less inferential and more intuitive response, though the method used made the answer prone to recency effects. In this experiment there was no prompt provided and participants may have chosen to refer more often to goals because these were the last and more recent individuals mentioned at the end of each first sentence.

As Crawley et al. (1990) suggested, pronoun reference resolution when answering questions is likely to be the result of interpreters' conscious reflections.[13] However, a forced-choice task like the one we developed could be hypothesized to trigger even less conscious, less inferential

---

13   See section 2.1.

responses than in Ferretti et al. (2009). Our results fit Stevenson et al.'s (1994) well-tested observation that end-states are the most salient component in the mental representation of events. As we have mentioned, Stevenson et al. (1994) found that in perfective contexts like (7) in section 2.1, in which subjects were sources, as in our task, participants generally preferred goals to sources.

This might be due to the fact that the verbs of transfer in the perfective contexts of our task involved completion of the relevant events. (26) entails, for example, that a meal was actually transferred, which can be taken to account for participants' predominant goal-oriented interpretations in the perfective condition. Conversely, (27) describes an action in progress, which can be taken to explain participants' preferred source-oriented readings in imperfective contexts.

(26)   Elizabeth took a meal to Mary. She breathed in the smell of fresh basil.
(27)   Elizabeth was taking a meal to Mary. She breathed in the smell of fresh basil.
        (adapted from Ferretti et al., 2009, Appendix A)

The fact that we found predominant goal-oriented readings in the perfective condition and source-oriented interpretations in imperfective contexts does not mean that our perfective targets could never lead to source-oriented interpretations or that our imperfective targets could never trigger goal-oriented resolutions. It merely means that all (or most) things being equal, perfective aspect is more likely to yield goal-oriented interpretations, while imperfective aspect is more likely to lead to source-oriented readings. Opposite/alternative readings can, indeed, still be achieved in conversation and were achieved to some extent by participants in our task. For example, in a context like (26), source-oriented interpretations might well be obtained if readers or hearers know that Elizabeth breathes in the smell of food whenever she serves a meal. We did not control for participants' social or cultural assumptions, personal experiences and/or the imaginary worlds that they could construct when assigning reference to our pronouns. Nonetheless, the fact remains that even when participants brought their own mental models to the task, perfective aspect yielded a statistically significant bias towards goals, while imperfective aspect triggered a statistically significant bias towards sources.[14] Furthermore, because the task tested the same targets in the perfective and the imperfective conditions and lexical meaning and grammatical structure in the second sentence of each target remained constant in both groups, it is clearly aspect that triggered more goal-oriented interpretations in targets like (26) and more source-oriented interpretations in targets like (27).

With regard to L2 learners, results partly overlapped with those obtained with the monolingual speakers: In the perfective condition, L2 speakers displayed a strong goal bias, consistent with what was found in monolinguals by us and by previous researchers (Kehler et al., 2008; Ferretti et al., 2009). In the imperfective condition, however, L2 learners differed from the monolinguals, and we found no significant bias either for sources or for goals.

As previously mentioned, German learners of English are an interesting group to use for comparison because aspect is not grammaticalized in German. Studies using German as the testing language indicate that German speakers may have an overall bias for end-states when processing new material in German (Carroll et al., 2008). As discussed above in relation to the study by Roberts and Liszka (2021), evidence suggests that these effects are also reflected in their L2 processing. Specifically, only the German learners displayed an overall end-state bias, whereas French learners did not.

We argue that the differences observed in our study between monolingual English speakers and German learners of English are to be attributed to this overall end-state bias German speakers may have incorporated in their general language processing, which may be extending to the processing of pronouns. For this reason, German learners of English may show a strong preference for goal interpretations in the perfective condition of our task—as a combined effect of aspect and baseline object bias—and no preference in the imperfective condition, since the overall object bias remains active, independently of the aspect, and it contrasts with the

---

14   We broadly use the term "mental model(s)" following Johnson-Laird (1983), Kintsch (1988), Morrow et al. (1987) and Frawley (1992), among others. In Frawley's (1992) terms, "We do not speak *about* the extensional world, but *about the model of the world we make up in our heads*" (emphasis in the original, p. 24).

subject bias typical of imperfective aspect. This finding contributes to the literature showing effects of L1 properties on L2 processing (Roberts et al., 2008; Jegerski et al., 2011; Contemori et al., 2019).

The potential confounding factor due to punctuation-induced conflicts within the experimental stimuli, as we discussed in relation to Roberts and Liszka's (2021) study, could be ruled out by the present experiment, as none of the target sentences included such conflicting information. Furthermore, whereas a smaller contribution of general L2 processing tendencies, such as suggested by the SSH, to the observed endpoint bias cannot be ruled out, previous research has already demonstrated that these domain-general aspects cannot be the only factors underlying these results. After all, the endpoint bias observed for L1 German learners of English was not present for L1 French learners in Roberts and Liszka's (2021) study.

7. LIMITATIONS

There are three limitations to this study that should be addressed by further research. One first issue is that the two groups of participants were tested with slightly different methodologies. While the group of English native speakers was tested in person, the group of L2 learners was tested online. Data collection started before the COVID pandemic and progressed during the period of restrictions, and as such we were forced to change the collection method. A second limitation lies in the metadata we collected. For L2 learners, we relied on their self-reported level of proficiency and we did not test them ourselves. Finally, our claim that the result with German learners is related to the transfer of an endpoint bias will only be fully reliable once it is shown that learners speaking a language without an endpoint bias display different preferences (for instance, it would be interesting to see if French learners of English display a source preference in imperfective aspect). Further research with participants of this kind is needed.

8. CONCLUSION

Through a forced-choice task, we reassessed findings from extant story continuation studies on monolinguals' pronominal reference assignments across sentences that contain English verbs of transfer in perfective and in imperfective aspect (Ferretti et al., 2009; Kehler et al., 2008) and extended this methodology to German learners of English. Like Ferretti et al. (2009) and Kehler et al. (2008), we found that in the monolinguals' group, goals were preferred in the perfective condition and sources were preferred in imperfective contexts. However, unlike Ferretti et al. (2009), we found no bias for goals in both aspects and, unlike Kehler et al. (2008), we found no overall bias for sources. Instead, we found a clear-cut bias towards goals in the perfective condition and sources in imperfective contexts. Additionally, by testing also German learners of English, we showed that while the pattern observed in monolingual speakers was replicated in the perfective condition, the effect was not found in imperfective contexts, possibly as a consequence of an overall end-state bias of German.

This study contributes to the body of research on anaphora resolution, and while on the one hand it reveals clear-cut findings in monolingual speakers, it extends them to L2 learners only partially. This is because it appears that L1 characteristics may affect anaphora resolution in L2 and mitigate the results. Furthermore, our study suggests that the tendency to try and identify only one factor affecting anaphora resolution, as illustrated in section 2.1, might need to be replaced by the study of a combination of features. For example, the results we obtained with L2 learners suggest that while aspect per se certainly has a role in anaphora resolution, the degree of grammaticalization of aspect in the L1 may have an additional, complementary, role.

APPENDICES

Appendices A and B are available at https://www.iris-database.org/details/37kht-lgvaS.

ETHICS AND CONSENT

This research was approved by the Research Ethics Committee of the University of Cambridge and by the Ethics Committee of Charles University, Prague. For the sake of transparency, raw anonymised data are made available at https://www.iris-database.org/details/KBUrw-Wg3Df.

FUNDING INFORMATION

AUTHOR AFFILIATIONS

Roberto B. Sileo orcid.org/0000-0002-2861-320X
University of Cambridge, UK

Luca Cilibrasi orcid.org/0000-0001-8262-0542
University of Siena, IT; Charles University, Prague, CZ

Julia Heine orcid.org/0000-0002-8781-5194
University of Cambridge, UK

Ianthi M. Tsimpli orcid.org/0000-0001-6015-7526
University of Cambridge, UK

REFERENCES

Andrews, A. (1985). The major functions of the noun phrase. In T. Shopen (Ed.), *Language typology and syntactic description: Clause structure* (Vol. I, pp. 132–223). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511619427.003

Baayen, R. H., Davidson, D. J., & **Bates, M. D.** (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. DOI: https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Sarkar, D., Bates, M. D., & **Matrix, L.** (2007). The lme4 package. *R package version, 2*(1), 74.

Brennan, S. E., Friedman, M. W., & **Pollard, C. J.** (1987). A centering approach to pronouns. *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics* (pp. 155–162). Association for Computational Linguistics. DOI: https://doi.org/10.3115/981175.981197

Broadbent, D. E. (1973). *In defence of empirical psychology*. Methuen. DOI: https://doi.org/10.4324/9781315797403

Buckledee, S. J. (2008). Universal Grammar and second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *New handbook of second language acquisition* (pp. 23–29). Brill Academic Publishers.

Carroll, M., Natale, S., & **Starren, M.** (2008). Acquisition du marquage du progressif par des apprenants germanophones de l'italien et néerlandophones du français [Acquisition of progressive marking by German learners of Italian and Dutch learners of French]. *Acquisition et Interaction en Langue Étrangère, 26*, 31–50. DOI: https://doi.org/10.4000/aile.2942

Clahsen, H., & **Felser, C.** (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3–42. DOI: https://doi.org/10.1017/S0142716406060024

Clahsen, H., & **Felser, C.** (2017). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition, 40*(3), 693–706. DOI: https://doi.org/10.1017/S0272263117000250

Contemori, C., Asiri, O., & **Irigoyen, E. D. P.** (2019). Anaphora resolution in L2 English: An analysis of discourse complexity and cross-linguistic interference. *Studies in Second Language Acquisition*, *41*(5), 971–998. DOI: https://doi.org/10.1017/S0272263119000111

Contemori, C., & **Dussias, P. E.** (2016). Referential choice in a second language: Evidence for a listener-oriented approach. *Language, Cognition and Neuroscience*, *31*(10), 1257–1272. DOI: https://doi.org/10.1080/23273798.2016.1220604

Crawley, R. A., Stevenson, R. J., & **Kleinman, D.** (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research, 19*(4), 245–264. DOI: https://doi.org/10.1007/BF01077259

Cunnings, I., Fotiadou, G., & **Tsimpli, I.** (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition*, *39*(4), 621–652. DOI: https://doi.org/10.1017/S0272263116000292

Depraetere, I. (1995). On the necessity of distinguishing between (un)boundedness and (a)telicity. *Linguistics and Philosophy, 18*, 1–19. DOI: https://doi.org/10.1007/BF00984959

Dosher, B. A. (2006). Working memory. *Encyclopedia of cognitive science*. DOI: https://doi.org/10.1002/0470018860.s00575

Ferretti, T. R., Rohde, H., Kehler, A., & **Crutchley, M.** (2009). Verb aspect, event structure, and coreferential processing. *Journal of Memory and Language, 61*, 191–205. DOI: https://doi.org/10.1016/j.jml.2009.04.001

Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). Holt, Rinehart, and Winston.

Frawley, M. (1992). *Linguistic Semantics*. Routledge. DOI: https://doi.org/10.4324/9781315044644

Frederiksen, J. R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes, 4*(4), 323–347. DOI: https://doi.org/10.1080/01638538109544525

Garvey, C., & **Caramazza, A.** (1974). Implicit causality in verbs. *Linguistic Inquiry, 5*(3), 459–464.

Garvey, C., Caramazza, A., & **Yates, J.** (1974–1975). Factors influencing assignment of pronoun antecedents. *Cognition, 3*(3), 227–243. DOI: https://doi.org/10.1016/0010-0277(74)90010-9

Gelormini-Lezama, C., Huepe, D., Herrera, E., Melloni, M., Manes, F., García, A. M., & **Ibáñez, A.** (2016). The overt pronoun constraint across three dialects of Spanish. *Journal of Psycholinguistic Research, 45*, 979–1000. DOI: https://doi.org/10.1007/s10936-016-9426-2

Grosz, B., Joshi, A. K., & **Weinstein, S.** (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics, 21*(2), 203–225.

Hirst, W., & **Brill, G. A.** (1980). Contextual aspects of pronoun assignment. *Journal of Verbal Learning and Verbal Behaviour, 19*, 168–175. DOI: https://doi.org/10.1016/S0022-5371(80)90152-8

Hobbs, J. R. (1976). *Pronoun Resolution*. City University of New York. DOI: https://doi.org/10.1145/1045283.1045292

Jegerski, J., VanPatten, B., & **Keating, G. D.** (2011). Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish. *Second Language Research, 27*(4), 481–507. DOI: https://doi.org/10.1177/0267658311406033

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.

Kanno, K. (1997). The acquisition of null and overt pronominals in Japanese by English speakers. *Second Language Research, 13*(3), 265–287. DOI: https://doi.org/10.1191/026765897673070746

Kehler, A., Kertz, L., Rohde, H., & **Elman, J. L.** (2008). Coherence and coreference revisited. *Journal of Semantics, 25*, 1–44. DOI: https://doi.org/10.1093/jos/ffm018

Kerkhofs, R., Vonk, W., Schriefers, H., & **Chwilla, D. J.** (2008). Sentence processing in the visual and auditory modality: Do comma and prosodic break have parallel functions? *Brain Research, 1224*, 102–118. DOI: https://doi.org/10.1016/j.brainres.2008.05.034

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182. DOI: https://doi.org/10.1037/0033-295X.95.2.163

Markov, I., Nastase, V., & **Strapparava, C.** (2018). Punctuation as native language interference. In E. M. Bender, L. Derczynski & P. Isabelle (Eds.), *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 3456–3466). Association for Computational Linguistics.

Moens, M., & **Steedman, M.** (1998). Temporal ontology and temporal reference. *Computational Linguistics, 14*, 15–28.

Montalbetti, M. M. (1984). *After binding: On the interpretation of pronouns* [Unpublished doctoral dissertation]. Massachusetts Institute of Technology.

Montrul, S., & **Slabakova, R.** (2003). Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in Second Language Acquisition, 25*(3), 351–398. DOI: https://doi.org/10.1017/S0272263103000159

Morrow, D. G., Greenspan, S. L., & **Bower, G. H.** (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26*, 165–187. DOI: https://doi.org/10.1016/0749-596X(87)90122-7

Pérez-Leroux, A. T., & **Glass, W. R.** (1999). Null anaphora in Spanish second language acquisition: Probabilistic versus generative approaches. *Second Language Research, 15*(2), 220–249. DOI: https://doi.org/10.1191/026765899676722648

Radford, A. (1988). *Transformational grammar: A first course*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511840425

Roberts, L., Gullberg, M., & **Indefrey, P.** (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition, 30*(3), 333–357. DOI: https://doi.org/10.1017/S0272263108080480

Roberts, L., & **Liszka, S. A.** (2021). Grammatical aspect and L2 learners' online processing of temporarily ambiguous sentences in English: A self-paced reading study with German, Dutch and French L2 learners. *Second Language Research, 37*(4), 619–647. DOI: https://doi.org/10.1177/0267658319895551

Rondal, J. A., Brédart, S., Leyen, N., Neuville, P., & **Pérée, F.** (1984). Coreference et strategie des fonctions paralleles dans le cas des pronoms anaphoriques ambigus. [Coreference and strategy of parallel functions in the case of ambiguous anaphoric pronouns]. *Cahiers de Psychologie Cognitive, 4*, 151–170.

Steinhauer, K., & **Friederici, A. D.** (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research, 30*(3), 267–295. DOI: https://doi.org/10.1023/A:1010443001646

Stevenson, R. J., Crawley, R. A., & **Kleinman, D.** (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes, 9*(4), 519–548. DOI: https://doi.org/10.1080/01690969408402130

Stoet, G. (2010). PsyToolkit – A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. DOI: https://doi.org/10.3758/BRM.42.4.1096

Walker, M. A., Joshi, A. K., & **Prince, E. F.** (1998). Centering in naturally occurring discourse: An overview. In M. A. Walker, A. K. Joshi & E. F. Prince (Eds.), *Centering theory in discourse* (pp. 1–28). Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198236870.003.0001

WHITE ROSE
UNIVERSITY PRESS
Universities of Leeds, Sheffield & York