AMBIGUITY RESOLUTION IN THE PRONOMINAL DOMAIN IN CHILD ROMANIAN

Ioana Stoicescu[*], Maria Aurelia Cotfas[**]

Abstract: This paper explores the acquisition of antecedent preferences by typically-developing Romanian-speaking (pre)school children. It investigates the antecedent preferences for subject pronouns (*pro, el* 'he', *this* 'acesta') in the subjunctive complements of volitional verbs. The results indicate that children overwhelmingly associated *pro* with close antecedents. There was a slight preference for a close antecedent with the overt pronoun, while no preference was observed for the demonstrative pronoun. The children treated *pro* and the overt pronoun differently and were aware that each pronominal marked a distinct degree of accessibility, but the evidence pointed towards a delay in the acquisition of the exact degree of accessibility marked by the overt pronoun. The delay was even more marked for the acquisition of the demonstrative pronoun.

Keywords: pronominal reference, antecedent preferences, accessibility

1. Introduction

Research on antecedent preferences for pronominal subjects has revealed that both syntactic and pragmatic cues play a part. In the framework of Accessibility Theory (Ariel 1990), these cues signal to comprehenders the accessibility (memory activation) of potential antecedents. Referential expressions are selected by speakers as accessibility markers, guiding the addressee in the retrieval of the discourse entity which the referential expression in question designates.

For acquisition, the question is whether children are able to identify the factors relevant for the determination of pronominal reference and whether some cues are easier to acquire than others.

Previous research on the acquisition of Romanian (presented in more detail in Section 4) has investigated several ambiguity inducing configurations which support the various degrees of accessibility of potential antecedents in different ways. Teodorescu (2016) used syntactic cues relevant for information structure and topicality in the line of Carminati (2002), while Stoicescu and Cotfas (2015) tested the combined influences of pragmatic factors and linear distance. However, since it was previously suggested that distance might be highly relevant for the interpretation of overt pronouns in some Romance languages (Filiaci et al. 2013), it is necessary to isolate and investigate this factor in more depth. In addition, in order to test the predictions of Accessibility Theory, it is also necessary to look at a wider range of pronominal elements than previously investigated in Stoicescu and Cotfas (2015). This study explores the antecedent preferences of both high and intermediate accessibility markers (null, personal overt and demonstrative pronouns, respectively).

[*] University of Bucharest, iodu2004@yahoo.co.uk.
[**] University of Bucharest, maura_cotfas@yahoo.com.

Bucharest Working Papers in Linguistics XX, 1, 65-82, ISSN 2069-9239, E-ISSN 2392-8093

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

This paper explores the acquisition of antecedent preferences by typically-developing Romanian-speaking (pre)school children. The question addressed is whether Romanian-speaking children have an adult-like mapping of reference relations for the null, personal overt and demonstrative pronouns, observing the requirements of Accessibility Theory (Ariel 1990, 2006). More specifically, it looks at the interpretation assigned to subject pronouns (*pro*, *el* 'he', *this* 'acesta') in the subjunctive complements of volitional verbs. The paper is organised as follows. Section 2 presents the tenets of Accessibility Theory. Section 3 describes the antecedent preferences of pronominal subjects in adult Romanian. Section 4 reviews the findings of previous research on the acquisition of pronominal biases in child Romanian. Section 5 outlines the experiment, the results and their discussion. Section 6 concludes the paper.

2. Accessibility Theory

Accessibility Theory accounts for the way in which referring expressions (REs) are used and interpreted in discourse. Ariel (1990) argues that the selection and interpretation of REs is determined by cognitive constraints applied to information storage. Not all information is readily accessible for retrieval. For instance, encyclopedic knowledge is stored in the long-term memory, and is retrieved with more difficulty than information about the speech event, which is highly accessible. Accessibility thus refers to the degree of memory activation of a particular mental representation. Representations that are more deeply embedded in our memories and are harder to retrieve are associated with a low degree of activation or accessibility. Representations that are stored in the short-term memory are regarded as having a high degree of accessibility.

During conversation, speakers build discourse models containing discourse entities that are either new or identifiable based on the information gathered from previous stretches of discourse, the extra-linguistic context or encyclopedic knowledge. According to the model proposed by Ariel (1990), the identification of the discourse entities referred to by NPs is helped by the fact that their accessibility is linguistically encoded. Speakers select certain types of referring expressions in order to indicate how accessible the discourse entity is, and guide addressees in retrieving the right mental representation. The linguistic coding of accessibility is organised in a scale of accessibility, given in (1) (from Ariel 1990: 73). The scale orders various types of NPs relative to the degree of accessibility they signal. For instance, full names indicate the lowest degree of accessibility, hence recourse to mental representations that are deeply embedded in long-term memory, while zero elements encode the highest accessibility. In the latter case, the addressee is directed to look for an antecedent which is either positioned in the immediate linguistic context or very salient.

(1) Full name + modifier > full name > long definite description > short definite description > last name > first name > distal demonstrative + modifier > proximate demonstrative + modifier > distal demonstrative + NP > proximate demonstrative + NP > distal demonstrative (−NP) > proximate demonstrative

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

(−NP) > stressed pronouns + gesture > stressed pronoun > unstressed pronoun > cliticized pronoun > verbal person agreement markers > zero

Ariel divides the referring expressions listed in (1) into three categories: (i) low accessibility markers (proper names, definite descriptions), used if their antecedent is not "currently salient" (Ariel 1990: 17); (ii) intermediate accessibility markers (deictics and demonstrative expressions); (iii) high accessibility markers (pronouns and gaps). By computing the frequency of occurrence for the antecedents of the three types of accessibility markers in English, Ariel (1990: 18) found that the antecedents of pronouns generally appear in the same or the previous sentence, while the antecedents of demonstratives occur either in the previous sentence or, more remotely, in the same paragraph. To illustrate this observation, consider the nominal phrases underlined in example (2). The DP *the affair between Helen and Paul* is a long definite description whose referent has low accessibility. *The review* is a shorter definite description, and it is also the topic, hence the mental representation of its referent has higher accessibility. The personal pronoun *he* refers back to Paul, a recently mentioned, highly accessible antecedent.

(2) LORI: when you were reading **the review**, you talked about **the affair between Helen and Paul**, […] all that happened was,
LINDA: was a kiss. […]
LORI: **He** kissed her.
(Santa Barbara Corpus: 023, example (1) in Ariel 2006: 15)

Ariel argues that the criteria on which the scale of referential expressions relies have universal application. These criteria are: (i) informativity, (ii) rigidity, and (iii) degree of attenuation. For instance, a long definite description like *the man who killed my cat* is more informative than a shorter definite description (*the man*), and it will be used as a low accessibility marker. Rigid expressions (e.g. proper names) identify a single referent. Attenuation refers to the phonological size of an expression. Heavier or stressed expressions are considered lower accessibility markers.

Accessibility *per se* is in its turn influenced by several factors: (i) the distance between the referential expression and the antecedent, (ii) the salience of the discourse entity, and (iii) the competition between discourse entities (if more antecedents are potentially available, then their accessibility diminishes). For the purposes of this study, the most important factor is distance. Distance matters because it is translatable into time or a more remote mention of a discourse entity, and this leads to the lower activation of the respective mental representation. Distance is not only related to the number of words between the anaphoric expression and its antecedent but also to syntactic, textual and pragmatic boundaries. This means that relative distance can be calculated taking into account various types of boundaries/units, namely the clause, the paragraph or the episode, or simply cohesive discourse chunks. As for salience, Ariel includes among salient discourse entities the speaker and the hearer (vs. third persons), humans (vs. inanimates), discourse topics (vs. non-discourse topics).

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

Ariel acknowledges that the scale itself is not universal. While the general order of expressions cannot vary from one language to another, the degree of accessibility marked by an individual expression may vary cross-linguistically (Ariel 1990: 75-76). This predicts that a pronoun may mark a slightly higher degree of accessibility in one language than in another.

Accessibility Theory accounts for both the selection of certain referential expressions in production, as well as the process of antecedent identification in comprehension. If speakers avail themselves of accessibility marking to make things easier for addressees, then the latter, given their own knowledge of accessibility, should be able to pick up the cues given by speakers and look for antecedents accordingly.

Summing up, the general predictions that can be made based on the universal ordering of the accessibility scale for null, overt, and demonstrative pronouns are the following: (i) null pronouns mark the highest degree of accessibility and should co-refer with the most accessible antecedents; (ii) overt pronouns are relatively lower on the scale than *pro* – they should be able to co-refer with less accessible antecedents than *pro*; furthermore, being still markers of high accessibility, they might also co-refer with very accessible antecedents but they should do so at lower rates than *pro*; (iii) demonstratives are markers of intermediate accessibility and should co-refer with less accessible antecedents, and should do so more frequently than overt pronouns.

3. Antecedent biases for overt and null pronouns in adult Romanian

Cotfas (2012) tested sentences in which the accessibility of the antecedent was determined by its relative distance to the pronoun. She investigated compound clauses in which the first conjunct contained a subject NP and the second conjunct included a matrix clause with another NP subject and a volitional verb, followed by a complement in the subjunctive mood (3a-b). The subject of the subjunctive complement was expressed by a null or overt pronoun.

Romanian has generally been described as a VSO language, such that the base position of the subject is post-verbal (Spec VP in Dobrovie-Sorin 1994, Spec VP and a postverbal Spec AgrP in Cornilescu 1997 for double subject constructions, Spec vP in Alboiu 2002) and case is assigned by the verb which moves to Inflection. Consequently, pre-verbal subjects have been analyzed as occupying non-argumental positions, i.e. as Themes (Dobrovie-Sorin 1994) or Topics (Cornilescu 1997, Alboiu 2002)[1]. For the purposes of this paper, we take the pre-verbal subject to signal topicality, since it encodes information already present in the previous discourse.

When the pre-verbal overt (nominal or pronominal) subject is embedded in a subjunctive complement, the subjunctive complementizer *ca* becomes overt (3b). Romanian subjunctives selected by volitional verbs are independent, both temporally and

[1] However, Motapanyane (1994) discusses cases when preverbal subjects are neither topics nor foci, suggesting that Spec IP must also be taken as an A-position. Likewise, more recently, Giurgea and Remberger (2014) and Giurgea (2017) discuss the frequency of SV occurrences and the existence of pre-verbal subjects that cannot be viewed as topicalized or focused.

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

with respect to control properties. Consequently, they define their own governing category for the embedded subject, which is liable to Principle B (Chomsky 1981), and thus interpretatively free in its relevant domain. The latter is both able to co-refer with or be disjoint from the matrix subject antecedent, regardless of whether it is a null or overt pronoun (3a-b). It behaves unlike its Romance counterparts, which have clear obviative behaviour in such configurations – illustrated for French in (3c).

(3) a. Marius$_i$ a intrat în birou şi Nicolae$_j$ vrea ***pro***$_{i/j}$
Marius has come into office and Nicolae wants *pro*
să plece.
SBJV leave-3SG
'Marius has come into the office and Nicolae wants him to leave.'

b. Marius$_i$ a intrat în birou şi Nicolae$_j$ vrea ca **el**$_{i/j}$
Marius has come into office and Nicolae wants that he
să plece.
SBJV leave-3SG
'Marius has come into the office and Nicolae wants that he should leave.'

c. Pierre$_i$ veut qu' **il**$_{*i/k}$ gagne la compétition.
Pierre wants that he win-SBJV-3SG the competition
'Pierre wants that he should win the competition.'

Cotfas (2012) found that the null subject of the subjunctive complement is interpreted as co-referential with the closest antecedent (89% of the time), while the overt pronominal subject is interpreted as disjoint from the closest, and co-referent with the more distant antecedent (94% of the time). Given the fact that the two antecedents are equally prominent in terms of topicality, their accessibility varies only in terms of distance, with the second NP being more salient than the first NP. The null pronoun prefers the closest most accessible antecedent, while the overt pronoun prefers the more distant one.

However, in a real-time experiment on the acquisition of pronominal reference testing sentences like (3a-b), the Romanian-speaking adult controls were at chance with respect to the antecedent of the overt pronoun and displayed no preference for the less accessible NP (Stoicescu and Cotfas 2015). What's more, when the pronouns are placed in complement clauses, with a single antecedent in the matrix (4), the overt pronoun also accepts the subject antecedent (66% of the time) (Pagurschi 2010).

(4) Ion crede că *pro* / el e inteligent.
Ion thinks that *pro* / he is intelligent
'John thinks that he is intelligent.'

The Romanian adult data point to the effects of the task used. The results in Cotfas (2012) are symmetrical, and might have been influenced by the nature of the task – the study used questionnaires administered in a written form, and the adult respondents had plenty of time to think, being induced to compare the two structures and assign distinct,

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

clear-cut interpretations to the two pronominal items. This is not what happens in real-time processing, when comprehenders have to give an optimal and quick response.

Cotfas (2012) claimed that her findings were in agreement with Accessibility Theory because it was assumed that the overt pronoun should mark lower accessibility and it is then only natural for it to be associated with a less salient antecedent (a similar claim was made in Stoicescu and Cotfas 2015). However, according to the accessibility scale, both the null and the overt pronouns are high accessibility markers, even though one is relatively higher than the other. Thus there should be more instances where *el* 'he' co-refers with the closest antecedent. And this is what we find in real-time experiments (Stoicescu and Cotfas 2015, Teodorescu 2016).

4. Previous findings on the acquisition of anaphora resolution in Romanian

Previous research on the acquisition of anaphoric relations has found that children are more permissive in their interpretation of pronouns than adults are. Children connect pronouns to antecedents in configurations where this would not be allowed in the adult grammar.

Studies on child Romanian have focused on the determination of reference in ambiguous contexts involving compound and complex clauses. When establishing antecedent preferences, Romanian-speaking children around the age of five appear to be sensitive to several factors, namely topicality, distance, and discourse relevance.

Stoicescu and Cotfas (2015) investigated the interpretation of *el* 'he', and *pro* by 5-year-old typically developing monolingual children, using compound clauses like the ones in (5) (similar to those used by Cotfas 2012), which summarized commissive (*I will do X*) or directive (*Do X*) scenarios. As topicality was counterbalanced in the test item, the only factors relevant for accessibility were distance and the discourse pragmatic influence of the speech acts.

(5) Introduction: The horse is the garden and the dinosaur comes too. They find an apple.
Dinosaur: "Hey, Horse, eat this apple!" (directive scenario)
Dinosaur: "I'll eat this apple!" (commissive scenario)

Test items*:*

a. [Calul$_i$ e în grădină] şi dinozaurul$_j$ vrea [***pro***$_{i/j}$ *pro*
horse-the is in garden and dinosaur-the wants
să mănânce mărul].
SBJV eat-3SG apple-the
'The horse is in the garden and the dinosaur wants to eat the apple.'

b. [Calul$_i$ e în grădină] şi dinozaurul$_j$ vrea [ca ***el***$_{i/j}$ *he*
horse-the is in garden and dinosaur-the wants that
să mănânce mărul].
SBJV eat-3SG apple-the
'The horse is in the garden and the dinosaur wants that *he* should eat the apple.'

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

The bias of the null pronoun for the most accessible antecedent, namely the second DP, was very clear only in the commissive condition (78%). In the directive condition, *pro* selected a close antecedent only 58% of the time, and was able to select a remote antecedent 21% of the time. As far as the overt pronoun is concerned, there was no visible bias for a more remote, less accessible antecedent. In both the commissive and directive scenarios, the children were basically at chance, displaying only a very marginal preference for the less accessible antecedent in the directive scenario and for the closest antecedent in the commissive scenario. The interaction of the scenario type with the antecedent preferences suggested that the children's antecedent preferences were guided by relevance considerations (Sperber and Wilson 1986), which seemed to override the distance factor.

One drawback of this study was that its design allowed for the children's potential *yes* bias to become manifest. The conditions in which the expected on-target response was acceptance of the test item were the conditions in which the children performed better. The speech acts in the introductory discourse interacted with the children's *yes* bias, and artificially improved their performance for the disjoint reference interpretation of the overt pronoun lowering their scores for the coreferential interpretation of *el* 'he'. Thus children were induced to accept both coreferential and disjoint reference readings for *el* 'he', and no clear preference pattern could be established. In the study presented here, the commissive or directive speech acts were excluded from the introduction, in order to investigate only the influence of distance and eliminate any *yes* bias effects.

Teodorescu (2016) explored the impact of topicality and the relative prominence of the syntactic position for the identification of the antecedents of pronominal subjects, starting from the Position of Antecedent Hypothesis proposed by Carminati (2002). She used test sentences in which the pronouns *pro*, *el* 'he' and the demonstrative *acesta* 'this' were part of a temporal adjunct, while the matrix included two possible antecedents: a subject DP and an object DP as in (6). The experiment used was a picture selection task.

(6) Elefantul a stropit motanul în timp ce ***pro*** / **el** / **acesta**
elephant-the has splashed cat-the in time what *pro* / he / this
mergea cu bicicleta.
went with bicycle-the
'The elephant splashed the cat while *pro/* he/ this was riding the bicycle.'

The Romanian children (3;11 – 5;11) showed no antecedent preferences in this task for the null and overt pronouns, and only a slight direct object preference in the demonstrative condition. Teodorescu explains the children's delay through a failure to turn to account syntactic and pragmatic information. The adults tested showed a preference for the subject antecedent with respect to the null pronoun and for the object pronoun with the demonstrative, but no preference with the overt pronoun. The adults in this study only partially confirmed the predictions of Carminati's hypothesis.

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

5. The experiment

5.1 Aim and predictions

This study aims to establish how typically-developing Romanian-speaking monolingual children aged 5-7 interpret *pro*, the unstressed personal pronoun *el* 'he', and the demonstrative *acesta* 'this', functioning as subjects of embedded subjunctive clauses selected by volitional verbs. More specifically, it investigates whether children are aware of the discourse interface constraints on the identification of antecedents, and whether they are sensitive to indicators of prominence/accessibility like closeness. Another question that arises is whether children know that pronominal elements in their language are ordered on an accessibility scale, and whether they have identified the degree of accessibility marked by each pronoun in their language.

The referential expressions tested mark different degrees of antecedent accessibility. If children are sensitive to the accessibility ordering of pronominals, we should see distinct response patterns for each individual pronominal tested. Consider the examples in (7). According to Accessibility Theory, the most accessible antecedent in such sentences is the matrix subject of the subjunctive clause; the first subject is less accessible, being more distant.

(7) a. Ursul$_i$ e la mare şi Remy$_j$ vrea ***pro***$_{i/j}$ să stea
bear-the is at seaside and Remy wants *pro* SBJV stay-3SG
la soare.
at sun
'The bear is at the seaside and Remy wants to lie in the sun.'

b. Ursul$_i$ e la mare şi Remy$_j$ vrea ca **el**$_{i/j}$ să stea
bear-the is at seaside and Remy wants that he SBJV lie-3SG
la soare.
at sun
'The bear is at the seaside and Remy wants that he should lie in the sun.'

c. Ursul$_i$ e la mare şi Remy$_j$ vrea ca **acesta**$_i$ să stea
bear-the is at seaside and Remy wants that this SBJV stay-3SG
la soare.
at sun
'The bear is at the seaside and Remy wants that this should lie in the sun.'

5.1.1 *pro*

According to the scale proposed by Ariel (1990), *pro* marks the highest degree of accessibility, so it should retrieve the most accessible antecedent. In (7a) *pro* should prefer the subject in the matrix of the subjunctive clause, the proper noun *Remy*, although, in principle, it could co-refer with the more distant subject DP *ursul* 'the bear' as well. If

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

children know that *pro* marks the highest accessibility, there should be a majority of responses where the closest antecedent is selected.

5.1.2 *el* 'he'

El 'he' is also a high accessibility marker, so it should be able to select prominent antecedents as well, but it also encodes lower accessibility than *pro*, so it is possible to link it to antecedents of relatively lower accessibility as well. In the complex clause in (7b), the personal pronoun can either co-refer with the matrix subject *Remy* or be coreferential with the more distant subject *ursul* 'the bear'. Since personal pronouns are lower on the scale than *pro*, they should retrieve a close antecedent less frequently than *pro*. However, since it is unclear how much lower than *pro* the position of *el* 'he' is on the scale, it is not easy to predict which antecedent will be favoured. According to Ariel's analysis of frequency conducted for English (1990: 18), pronouns are primarily connected to antecedents in the previous sentence (60% of the time), and, only secondarily to antecedents in the same sentence (20%). However, these results are reported for English, which is a non-pro-drop language. The Romanian adults tested by Cotfas (2012) preferred the more remote antecedent as well. This preference is also likely to occur in Romanian for considerations related to Grice's Maxim of Quantity (Grice 1975) – there is already a mechanism for the retrieval of very close antecedents, namely the null pronoun *pro*. These are arguments for a remote antecedent preference for *el* 'he', with the caveat that close antecedent responses can also occur. If children know the ordering scale and distinguish *el* 'he' from *pro*, we should see a different response pattern from the null pronoun condition (less close antecedent responses than in the null pronoun condition, and a preference for the remote antecedent).

5.1.3 *acesta* 'this'

The demonstrative pronoun *acesta* 'this' is an intermediate accessibility marker. According to Ariel, when used anaphorically, demonstratives signal that the mental representation of the referent is more deeply embedded in memory. Demonstratives resemble personal pronouns in that they are associated with more distant antecedents, but, unlike personal pronouns, demonstratives do not select highly accessible antecedents. Ariel points out that, in the English texts she analysed, both demonstratives and personal pronouns were preferentially linked to antecedents in a previous sentence (at the same rate, around 60%), but they differed in their secondary preferred antecedent position (pronouns – same sentence, demonstratives – more remotely, in the same paragraph (20%). Consequently, the prediction that can be made for the Romanian demonstrative *acesta* 'this' is that it should be associated with less accessible antecedents than *pro*. Indeed, in (7c) *acesta* 'this' cannot corefer with the matrix subject *Remy*, which is too accessible. It can only be coreferential with the less accessible antecedent, namely the subject of the first conjunct, *ursul* 'the bear'.

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

To sum up, if Romanian-speaking children are aware of the accessibility scale, a) they should prefer the closest DP as an antecedent for *pro*; b) they should prefer the more distant DP as an antecedent for *el* 'he', but they should also select the closer DP even if they do so less frequently than in the case of *pro*; c) they should prefer the remote DP as the antecedent for *acesta* 'this'. The predicted preferences are summarised in (8):

(8) Selection of antecedents predicted by Accessibility Theory
  (i) *pro*: close DP
  (ii) *el* 'he': remote DP (preferred) and close DP (at lower rates than *pro*)
  (iii) *acesta* 'this': remote DP

5.2 Participants

The experiment was administered to three groups of Romanian-speaking typically-developing children: (i) twenty 5 year-olds (mean age 5;5, age range 4;10-5;9, SD = 3.4); (ii) eighteen 6 year-olds (mean age 6;3, age range 5;10-7;1, SD = 4); (iii) twenty-three 7-year-olds (mean age 7;7, age range 6;8-8;7, SD = 5.6). They were recruited from a kindergarten and a primary school in Bucharest. All children were tested individually, in a single session. A control group of eight adults was also tested.

5.3 Procedure, materials and design

The experiment consisted of a binary judgment task that followed a short introduction acted out with props by the experimenters. The task was the same as the one used in Stoicescu and Cotfas (2015), but the directive/commissive speech acts were eliminated. The participants were told that they were going to hear a story about two characters. Then they were introduced to a hand-puppet, Grandma, who told them what happened next in the story. The experimenters showed the children the toys representing the two characters, and acted out the beginning of the story. Grandma completed the story with a single compound clause, which represented the test item (9-11). Finally, the experimenter asked two *wh*-questions: *Who is going to...?* and *What does X want to happen?* The purpose of the first question was to elicit the name of the character that was regarded as the antecedent of the pronominal tested. The second question checked comprehension and confirmed that the referential relations were indeed the ones indicated by the answer to the first question (children's responses to the clarification question generally included (at least) one overt DP and truncated responses with a null subject and a subjunctive verb of the type *să stea la soare* "that (he) should sunbathe" were rather scarce, so it was not hard to establish the antecedent that the child had considered). There were three experimental conditions corresponding to *pro*, the personal pronoun *el* 'he', and the demonstrative *acesta* 'this' (9-11). There were three practice stories (one per condition), four test items per condition and four fillers. The list of items was randomized by condition and potential response type.

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

(9) Condition 1: *pro*

Ursul e la mare şi vine şi Remy şi discută despre statul la plajă. Bunica: **Ursul e la mare şi Remy vrea *pro* să stea la soare.** Experimenter: Cine o să stea la soare? Ce vrea Remy să se întâmple?	The bear is at the seaside and Remy [a mouse] comes along too and they talk about sunbathing. Grandma: **The bear is at the seaside and Remy wants to sunbathe.** Experimenter: Who is going to sunbathe? What does Remy want to happen?

(10) Condition 2: *el* 'he'

Şoarecele Remy este în livada cu cireşi şi se iveşte şi porcul şi încep să vorbească despre cireşe. Bunica: **Remy e în livadă şi porcul vrea ca el să culeagă cireşele.** Experimenter: Cine o să culeagă cireşele? Ce vrea porcul să se întâmple?	Remy the mouse is in the cherry orchard. The pig comes along too. They start talking about cherries. Grandma: **The mouse is in the orchard and the pig wants that he should pick the cherries.** Experimenter: Who is going to pick the cherries? What does the pig want to happen?

(11) Condition 3: *acesta* 'this'

Piratul intră într-un magazin. Vine şi soldatul şi văd un telefon mare! Bunica: **Piratul a venit la magazin şi soldatul vrea ca acesta să cumpere telefonul.** Experimenter: Cine o să cumpere telefonul? Ce vrea soldatul să se întâmple?	The pirate enters a store. The soldier comes along too and they see a big phone. Grandma: **The pirate came to the store and the soldier wants that this should buy the phone.** Experimenter: Who is going to buy the phone? What does the soldier want to happen?

During the experiment, care was taken not to stress the overt pronoun (to avoid contrastive focus, which would have triggered obligatory coreference between the pronoun and the matrix subject).

The structure of the filler items was similar to the one of the test items but without the ambiguity (e.g. *The bear bought some fishing lines and the dinosaur wants to fish*). The fillers included narrow focus questions with respect to the first or the second subject (*Who bought some fishing lines?* or *Who wants to fish?*). Their purpose was to check that the children were paying attention and to prevent them from building an irrelevant answering strategy. If a child responded more than once with the same type of antecedent (e.g. the close one), it was a sign that a certain response strategy was at work, so, in the filler, we would ask the question whose answer was the remote DP, in order to make the child consider an alternative antecedent (as recommended in Crain and Thornton 2000).

The first conjunct of the test sentence, a simple indicative clause, repeats the introduction to the story, and its truth is thus presupposed. The second conjunct gives new

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

information about the wishes of one of the characters. It is a complex clause with an indicative matrix, and a subjunctive complement. There are three subjects in the test item. The first one is the subject DP of the first conjunct, the second is the subject of the main clause in the second conjunct, and the third (the pronominal tested) is the subject of the subjunctive complement. All DPs have matching gender and number features. The test item was thus ambiguous providing two possible antecedents for the pronominal. Since it was noted that distance may be overridden by topicality (Ariel 1990: 19), the entire scenario was balanced with respect to topicality – each discourse entity was alternately the topic, and was mentioned an equal number of times. The only criterion which differentiates the two main subjects in terms of accessibility is distance. The scenarios were equally plausible for each character.

5.4 Results

The results are presented in Figures 1-4, which display the mean percentages of close and remote antecedent responses for each age group. The one-sample *t*-test was run to check for strong antecedent preferences. If the scores for a particular antecedent were significantly above chance, this indicated that the relevant position was favoured well and above the other.

All age groups chose the close antecedent for *pro* at levels significantly higher than chance, which proves a clear preference. The one-sample *t*-test showed significant differences from chance for the close antecedent responses in the *pro* condition for (i) 5-year-olds: $t(19) = 10.376$, $p < .001$); (ii) 6-year-olds: $t(17) = 6.985$, $p < .001$; (iii) 7-year-olds: $t(22) = 22.597$, $p < .001$. Adults were at ceiling in this condition.

In the *el* 'he' condition, no preference for the remote antecedent was visible with any child group as the relevant scores did not go significantly above chance ((i) 5-year-olds: $t(19) = -1.778$, $p = .09$); (ii) 6-year-olds: $t(17) = -2.06$, $p = .055$; (iii) 7-year-olds: $t(22) = -.397$ $p = .695$). Adults, however, were significantly above chance with respect to the remote antecedent ($t(7) = 1.922$, $p < .05$), which signals a clear preference for the remote antecedent for the overt pronoun.

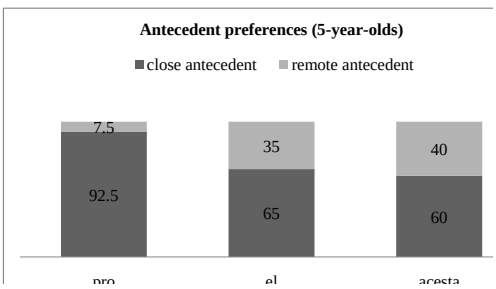


Figure 1. Antecedent preferences (5-year-olds)

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

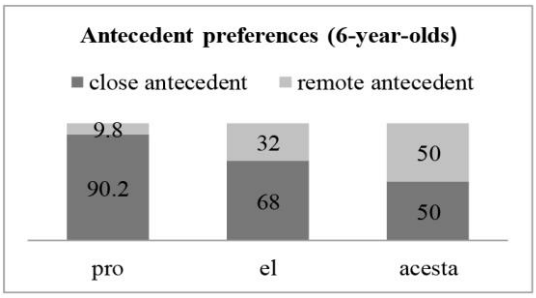


Figure 2. Antecedent preferences (6-year-olds)

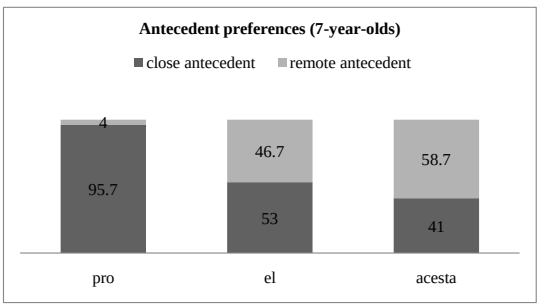


Figure 3. Antecedent preferences (7-year-olds)

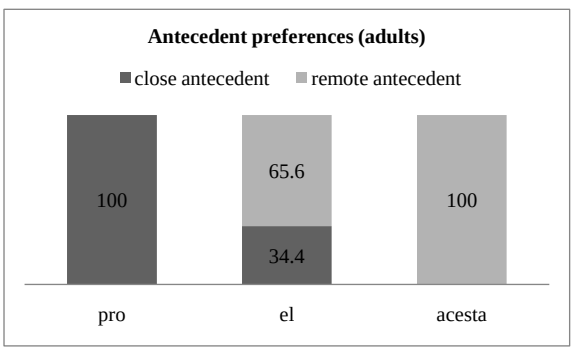


Figure 4. Antecedent preferences (adults)

We ran a separate analysis for the overt pronoun condition, considering the child responses that identified the closer DP as the antecedent of the personal pronoun. As reported in Figures 1 and 2, the 5 and 6-year-old children seemed to prefer the closer antecedent for *el* 'he'. In order to check the robustness of this preference, the one-sample *t*-test was used to identify whether the means for the close antecedent responses went significantly above chance. However the scores for this type of response fell short of statistical significance irrespective of age group: (i) 5-year-olds: $t(19) = 2.041$, $p = .055$; (ii) 6-year-olds: $t(17) = 2.060$, $p = .055$; (iii) 7-year-olds: $t(22) = .397$, $p = .695$. Statistically, the children were at chance in the overt pronoun condition.

In the demonstrative condition, with respect to the remote antecedent scores, all child groups were at chance (i) 5-year-olds: $t(19) = -1.140$, $p = .269$); (ii) 6-year-olds

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

were obviously at chance (50% for either antecedent); (iii) 7-year-olds: $t(22) = 1.073$, $p = .295$). Adults were at ceiling in this condition, they always chose a remote antecedent for the demonstrative. We ran a separate analysis for the 5-year-old group for close antecedent responses since these children seemed to display a 60% preference (see Figure 1). However, the scores for the close antecedent responses did not reach significance ($t(19) = 1.140$, $p = .269$). All children were at chance in this condition.

The child data were also analysed using one-way ANOVA, in order to check whether there were significant differences between the three groups of children. There were no statistically significant differences between the group means in any of the conditions tested. In the null pronoun condition, ANOVA showed no effect of age on performance (F (2,58) = .468, $p = .629$). In this condition, the target means are above 90% for all groups. No statistically significant difference was noted in the overt pronoun condition (F (2,58) = .887, $p = .418$). The children's target means ranged between 32% and 47%. The same happened in the demonstrative condition, where no significant differences were found (F (2,58) = 1.172, $p = .317$). For the demonstrative, the target means hovered around 50% for all child groups.

The independent samples *t*-test was used to check for differences between the three groups of children and the adult controls. The test was run on mean scores for both close / remote antecedent responses for *el* 'he' and *acesta* 'this', so as to check whether the child and adult results were skewed in the same direction. For the null pronoun, all age groups preferred the close antecedent, so the *t*-test was run only on the close antecedent responses.The test indicated that children and adults treated *pro* similarly (there were no significant differences between the adults and (i) the 6-year-olds ($t(24) = -1.686$, $p = .11$); (ii) the 5-year-olds ($t(26) = -1.831$, $p = .83$). There was a significant difference between adults and the 7-year-olds ($t(22) = -2.152$, $p < .05$) in this condition, but the mean percentages are about 90% for both age groups, so not much can be made of the statistical difference found.

In the overt pronoun condition, the results of the 7-year-old children did not differ from the results of the adults in a statistically significant way for either remote antecedent responses ($t(25.853) = -1.795$, $p = .084$) or close antecedent responses ($t(25.853) = 1.795$, $p = .084$). Yet, as can be seen in Figures 3 and 4, despite the absence of a statistical difference, it cannot be claimed that the 7-year-olds are adult-like with respect to *el* 'he', because their response pattern is the opposite of the adults' response pattern. In addition, adults performed significantly differently from the 6-year-olds (with respect to both remote antecedent responses ($t(24) = -2.412$, $p < .05$), and close antecedent responses ($t(24) = 2.412$, $p < .05$). Adults also performed significantly differently from the 5-year-olds (with respect to both remote antecedent responses ($t(24.459) = -2.873$, $p < .05$), and close antecedent responses ($t(23.758) = 3.086$, $p < .05$)).

In the demonstrative pronoun condition, none of the child groups were adult-like. The *t*-test revealed significant differences between adults and the 7-year-olds (with respect to both remote antecedent responses ($t(22) = -5.095$, $p < .001$, and close antecedent responses ($t(22) = 5.095$, $p < .001$); adults also performed significantly differently from the 6-year-olds (with respect to both remote antecedent responses ($t(17) = -5.05$, $p < 0.001$), and close antecedent responses ($t(17) = 5.05$, $p < 0.001$). Adults also performed significantly differently from the 5-year-olds (with respect to both

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

remote antecedent responses ($t(19) = -6.839$, $p < .001$), and close antecedent responses ($t(19) = 6.839$, $p < .001$)).

We also performed an analysis of the children's individual response patterns, which revealed that, generally, they did not resort to guessing. Table 1 presents the number of children corresponding to a certain response pattern in the overt pronoun condition. We regarded the instances in which a certain antecedent was chosen at least 75% of the times as evidence of a preference. As can be seen in Table 1, the younger children (the 5- and the 6-year-olds) favoured the close antecedent for *el* 'he'. The 7-year-olds seem to be divided with respect to their antecedent preferences.

Table 1. Number of children corresponding to particular response patterns in the overt pronoun condition

	5-year-olds	6-year-olds	7-year-olds
guessing pattern (50%)	3	1	4
preference for the remote antecedent	6	4	9
preference for the close antecedent	11	13	10

As reported in Table 2, in the demonstrative condition, the number of children who were guessing, being 50% correct, was small as well. The remaining children are approximately equally distributed – the youngest are slightly more inclined towards the close antecedent. Within the 6-year-old group, the children are almost equally divided between the two options. With the oldest children a trend towards selecting the more distant antecedent emerges.

Table 2. Number of children corresponding to particular response patterns in the demonstrative pronoun condition

	5-year-olds	6-year-olds	7-year-olds
guessing pattern (50%)		3	3
preference for the remote antecedent	8	8	13
preference for the close antecedent	12	7	7

Summing up the results, the main findings are the following: firstly, children performed at adult levels with respect to the null pronoun. For *el* 'he' and *acesta* 'this', the children's responses did not go significantly above chance for any of the three groups. The children's performance did not vary significantly with age for any of the pronouns tested. However, by looking at the children's individual responses, we noted that there were response patterns in the overt pronoun condition, in which a big percentage of the children (especially the 5- and 6-year-olds) had a preference for the proximate antecedent, while the 7-year-olds had expanded their choices to include a remote antecedent as well. For the demonstrative, the 5-year-olds went with the close antecedent; the 6-year-olds were equally divided, while the 7-year-olds progressed towards the remote antecedent choice.

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

5.5 Discussion

Let us start with a discussion of the adult response patterns which reveal something about the accessibility scale instantiated in Romanian. The adult responses conformed to the pattern predicted by Accessibility Theory. In the null and demonstrative pronoun conditions, adults displayed strong preferences. Since adults connected *pro* with the closest antecedent available, and the demonstrative *acesta* 'this' with a remote antecedent, it is clear that there is a contrast between the degrees of accessibility marked by these pronouns, which confirms that they should be placed in different categories (*acesta* 'this' as an intermediate accessibility marker, and *pro* as a high accessibility marker). The predictions made were confirmed for the personal pronoun *el* 'he' as well. There was a statistically significant preference for the more remote antecedent (65.6%), but there were also close antecedent responses (but at a lower rate than in the *pro* condition – 49 % vs. 97%). This proves that *el* 'he' is a high accessibility marker in Romanian, selecting two potential antecedents in the linguistic context – either within a short distance or more remotely. The fact that the closer DP was selected less frequently in the overt pronoun condition than in the *pro* condition shows that, while still being high accessibility markers, overt pronouns are lower on the scale than *pro*. In Cotfas (2012), the preference for a distant antecedent with *el* 'he' was more marked than in this experiment because of the experimental method used.

Moving on to the children's results, all age groups performed at adult rates in the null pronominal condition, where the closer antecedent was favoured. This preference was very visible for all the participants. In the *el* 'he' condition, the target means themselves and the response patterns suggested that the children had a slight preference for the closer antecedent, even though it did not reach significance. This preference was more marked in the younger groups, at ages 5 and 6 but it was not a very strong one, unlike in the case of *pro*. However, it cannot be argued that the children treated the overt and the null pronouns the same, since there were distinct response patterns (with more variation in the performance for *el* 'he' than *pro*). However, the 7-year-old children were closer to the adult norm, a sign that they were starting to realise that the overt pronoun plays a different role in discourse and can pick more distant antecedents as well.

The analysis of individual responses showed that, generally, the children did not resort to guessing for *el* 'he'. We propose that the preference for the closer antecedent is the result of relevance considerations, namely the Principle of Optimal Relevance (Sperber and Wilson 1986, Wilson 1992) and a tendency to save on processing resources by building local representations (Ferreira and Patson 2007). According to Relevance Theory (Wilson 1992: 175), an interpretation is optimally relevant iff: (i) it is relevant enough for the hearer to pay attention to it; (ii) it is economical in terms of processing effort. Obviously, the selection of an interpretation in which the pronoun was associated with the closest antecedent meets these relevance criteria. The choice of the closest antecedent puts less pressure on working memory.

Moreover, Ferreira and Patson (2007) argued that in fast communication, the parser tends to compute locally rather than globally, building "good enough" representations that can be revised during conversation. These considerations explain the children's tendency to favour coreference with a local, proximate antecedent.

BDD-A29082 © 2018 Universitatea din București
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

A reason for the poor results of the children in the demonstrative condition is their limited exposure to this pronoun used anaphorically. *Acesta* 'this' is mostly used in the formal register and is highly infrequent in the longitudinal or cross-sectional corpora of child language. In the corpus of frog story narratives compiled by Buja (2008) (comprising data collected from thirty-seven children (age range 3;2-9;11) and ten adults), we found four instances of *acesta* 'this' in child speech and four instances in adult speech. Other demonstratives that belong to the spoken register (*ăsta* 'this', *ăla* 'that') were used deictically in this corpus. Moreover, we could not find *acesta* 'this' in the longitudinal corpus belonging to child Iosif (Stoicescu 2013) (covering the age span 1;10-3;1) (although informal demonstratives like *ăsta* 'this', *ăla* 'that' were present, being used deictically) .

Summing up, in child Romanian, at the ages tested, the referential options for *pro* are available – the children know that *pro* has to corefer with a highly accessible antecedent, in accordance with Accessibility Theory. However, children do not know the exact degree of accessibility marked by *el* 'he' and *acesta* 'this' in adult Romanian, although there is evidence that they start resorting less to the local computation and move towards the adult norm.

6. Conclusions

This study confirmed that Accessibility Theory is a mechanism used by adult Romanian speakers. The adults tested distinguished between *pro*, *el*, and *acesta*, according to the predictions of Accessibility Theory. Thus *pro* was connected to a close highly accessible antecedent, while the overt pronoun, a high accessibility marker as well, retrieved both close and more distant antecedents, with a slight preference for the latter. However, the demonstrative, an intermediate accessibility marker, was only associated with a distant antecedent. Romanian children know that referential expressions are linked to different types of antecedents, as evidenced by their distinct response patterns for each pronoun. They are aware of the general principle of Accessibility Theory but they learn the degrees of accessibility marked by each type of linguistic expressions gradually – at the age of 5 they know that *pro* has a highly accessible antecedent, but they do not know the accessibility status of other expressions like *el* 'he' and *acesta* 'this'. In order to interpret such elements, Romanian-speaking children initially use a default mechanism – building a representation that is efficient processing-wise, based on the computation of local environments. This generated the preference for the closer antecedent for overt pronominals. There is also evidence of progress towards the adult norm at the age of seven.

References

Alboiu, G. 2002. *The Features of Movement in Romanian*. Bucharest: Editura Universităţii din Bucureşti.
Ariel, M. 1990. *Accessing Noun-Phrase Antecedents.* London: Routledge.
Ariel, M. 2006. Accessibility Theory. In K. Brown, A. H. Anderson, L. Bauer, M. Berns, G. Hirst, J. Miller (eds.), *Encyclopedia of Language and Linguistics*, vol. 1, 15-18. Amsterdam: Elsevier.

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)

Avrutin, S. and Wexler, K. 2000. Children's knowledge of subjunctive clauses: Obviation, binding and reference. *Language acquisition* 1 (8): 69-102.

Buja, E. 2008. *Relating Events in Narrative: A Case Study of Romanian*. Braşov: Editura Universităţii Transilvania.

Carminati, M. N. 2002. The Processing of Italian Subject Pronouns. PhD dissertation, University of Massachusetts, Amherst.

Chien, Y., Wexler, K. 1990 Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1 (3): 225-295.

Chomsky, N. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Berlin: Mouton de Gruyter.

Cornilescu, A. 1997. The double subject construction in Romanian. Notes on the syntax of the subject. *Revue Roumaine de Linguistique* XLII (3-4): 101-147.

Cotfas, M. A. 2012. On the Syntax of the Romanian Subjunctive. PhD dissertation, University of Bucharest.

Cotfas, M.A. 2017. On the left periphery of independent subjunctives in Romanian: Topics, foci and complementizer deletion. *Revue roumaine de linguistique* LXII (3): 253-278.

Crain, S., Thornton, R. 2000. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press.

Dobrovie-Sorin, C. 1994. *The Syntax of Romanian*. Berlin · New York: Mouton de Gruyter.

Ferreira, F. and Patson, N.D. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1 (1-2): 71-83.

Filiaci, F., Sorace, A. and Carreiras, M. 2013. Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language and Cognitive Processes* 29 (7): 825-843.

Giurgea, I., Remberger, E. 2014. Polarity fronting in Romanian and Sardinian. In M.-H. Côté, E. Mathieu (eds.), *Variation within and across Romance Languages. Selected Papers from the 41st Linguistic Symposium on Romance Languages (LSRL), Ottawa, 5-7 May 2011*, 173-197. Amsterdam/ Philadelphia: John Benjamins.

Giurgea, I. 2017. Preverbal subjects and topic marking in Romanian. *Revue roumaine de linguistique* LXII (3): 279-322.

Grice, P. 1975. Logic and conversation. In P. Cole, J. Morgan (eds.), *Syntax and Semantics*, vol 3, *Speech Acts*, 41-58. New York: Academic Press.

Motapanyane, V. 1994. On preverbal positions in Romanian. *Revue canadienne de linguistique / Canadian Journal of Linguistics* 39: 15-36.

Pagurschi, F. 2010. Interpreting Pronouns in Discourse Representation Theory. PhD dissertation, University of Bucharest.

Silva, C. 2013. Acquiring pronominal subjects in European Portuguese: Interpretation in indicative and subjunctive contexts. Paper presented at the 11th Generative Approaches to Language Acquisition Conference, 5-7 September 2013, University of Oldenburg.

Sperber, D. and Wilson, D. 1986. *Relevance*. Oxford: Blackwell.

Stoicescu, I. 2013. *The Acquisition of Tense and Aspect in Romanian*. Bucharest: Editura Universităţii din Bucureşti.

Stoicescu I. and Cotfas, M. 2015.Testing Romanian-speaking children's knowledge of Accessibility Theory in subjunctive complements. Paper presented at the 13th Conference on British and American Studies, 8-9 May, University of Braşov.

Teodorescu, O. 2016. On pronominal anaphora resolution in temporal adjuncts in child Romanian. *Bucharest Working Papers in Linguistics* 13 (1): 37-48.

Wilson, D. 1992. Reference and relevance. *UCL Working Papers in Linguistics* 4: 165-191.

BDD-A29082 © 2018 Universitatea din Bucureşti
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 12:53:24 UTC)