Private and inner speech and the regulation of social speech communication

Conchi San Martín Martínez a,∗, Humbert Boada i Calbet a, Peter Feigenbaum b

a Facultat de Psicologia, Universitat de Barcelona, Spain
b Fordham University, United States

A R T I C L E   I N F O

Keywords:
Private speech
Inner speech
Referential communication
Discourse
Vygotsky

A B S T R A C T

To further investigate the possible regulatory role of private and inner speech in the context of referential social speech communications, a set of clear and systematically applied measures is needed. This study addresses this need by introducing a rigorous method for identifying private speech and certain sharply defined instances of inaudible inner speech. Using this classification system, longitudinal data were gathered from 10 pairs of children performing a referential communication task at 4.5, 6.5, and 8.5 years of age. Results demonstrated children’s substantial production of private and inner speech in this communicative situation, with speech forms varying in amount and type as a function of age, communicative role (speaker or listener), and the complexity of the material to be communicated. It is suggested that private and inner speech embedded in discourse may serve a regulatory role in social speech communication.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Vygotsky (1934/1987) claimed that children talk aloud to themselves when they encounter a cognitive challenge that requires them to stop and think. This personal form of verbal thinking, called private speech, helps children to analyze and solve problems and gradually develops into a silent, subvocalized form of thinking called inner speech. Studies of private speech development (see Berk, 1992;
Winsler, Fernyhough, & Montero, 2009, for reviews) have focused largely on problem-solving tasks of an academic nature, but recently researchers working with referential communication tasks (see Martin, Boada, & Forns, 2009, for review) have begun to explore the role of private and inner speech in solving problems related to social speech communication. One particularly promising avenue of investigation taken by Boada and Forns (2004), for example, uses a communication task to examine the relations between private speech and children’s personal understanding and socially communicative use of abstract spatial deictic terms. These pioneering studies have made strides in documenting the presence of private speech in speakers’ messages and listeners’ actions, and in tracking the developmental transformation of private speech into inner speech, as well as linking the cognitive functions of private speech to task difficulty. However, further progress is greatly hindered by methodological obstacles. Specifically, the use of different classification systems for identifying and categorizing private speech and inner speech has made it very difficult to compare results across tasks.

The present study addressed these issues by introducing a rigorous methodology for identifying private speech and certain sharply defined instances of inaudible inner speech. We designed an empirical method that could reliably distinguish between utterances intended as social, private, and inner speech, and we applied this system to data collected from 10 pairs of children examined longitudinally at 4.5, 6.5, and 8.5 years of age. Our aims were to replicate previous findings showing that children produce private and inner speech in communicative situations and to provide a more in-depth examination of how these speech forms vary in amount and function according to age, communicative role, and the cognitive complexity of the material to be communicated.

1.1. Eliciting private speech in a communication task

Because communicative exchange with speech requires the coordination of a host of cognitive, linguistic, and social abilities (Lefebvre-Pinard, 1985), a connection between private speech and the development of socially adapted speech would be expected (Wertsch, 1985). Among the communicative tasks that present a major cognitive challenge for speech regulation are referential tasks. In these tasks, the speaker has to select and verbally codify in a message an object or action, as well as convey information about its relationships, such as explaining how different objects are situated in a picture or giving instructions to follow a route. The message transmitted to the listener must unambiguously describe the attributes that distinguish one referent from another. Speaker and listener are prevented from seeing one another; thus, they must rely entirely on verbal communication to coordinate their efforts (Dickson, 1981). There is general agreement that during the early school years children show significant progress in the regulation of their communications, especially at about 6 years of age, although as listeners they still have difficulty detecting ambiguity in the information that is received in conversation, as well as difficulty producing unambiguous messages as speakers before 8–10 years of age (Boada & Forns, 1997; Camaioni, Ercolani, & Lloyd, 1998; Lloyd, Mann, & Peers, 1998).

Typically, private speech has not been included in most referential communication studies as a possible indicator of young children’s intentions to regulate their speech. Thus, the few studies that have explicitly taken up that question (Boada & Forns, 2004; Girbau, 1997; Varenne & Beaudichon, 1996) are breaking new ground. Before considering them, we address the critical question of the empirical measures that enable researchers to reliably identify utterances of private speech occurring in the middle of a social speech communication.

1.2. Empirical methods for studying private speech in communicative tasks

Classical methods of distinguishing utterances of private speech from those of social speech, which were developed for use in non-communicative or minimally communicative tasks, are based on the notion that private speech is not directed to another person, but rather to oneself (Winsler, Fernyhough, McClaren, & Way, 2005). Therefore, one set of “positive” indicators of private speech includes behaviors such as lowered voice volume, lack of eye contact, and lack of physical contact. These behaviors contrast sharply with those associated with social speech communication, which is typically signaled by direct eye contact or visual gaze directed toward the listener (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1977), the use of vocatives (Dickie, 1973), body oriented toward the
listener, “broadcasting” speech if the listener is not in visual contact (Furrow, 1980), and other markers of social engagement. Thus, another set of indicators that has been used to identify private speech is the complete absence of all social speech indicators (Feigenbaum, 1992).

In the context of social or speech communication, however, determining whether an utterance is really directed to another person or to the self is especially difficult (Diaz, 1992) and crucial to accomplish (San Martín, Boada, & Forns, 2009), suggesting that indicators other than the classical ones are needed. Dysfluencies in speech – such as repetitions, self-corrections, hesitation pauses, and even long silences – are common and may be associated with either private speech or social speech. Consequently, researchers have struggled to develop more specific methods of identifying private elements occurring within communicative contexts (Boada & Forns, 2004; Frawley & Lantolf, 1985; Girbau, 1997, 2002; Lantolf, 2003; Varenne & Beaudichon, 1996). In the area of referential communication tasks, Varenne and Beaudichon (1996) base the identification of private speech exclusively on the presence of self-corrections in the messages, while Boada and Forns (2004) infer private speech solely by the presence of certain types of fragmentations in the messages. Girbau (1997) infers private speech using some of the positive indicators of private speech. Although the cited studies employ different category systems, it has become more evident that the emergence of the ability to regulate referential communication using private speech occurs in the early school years – at about age 6, being somewhat less pronounced at age 4 (Boada & Forns, 2004). This speech has also been documented in cross-sectional studies in the age range of 8–10 (Girbau, 1997; Varenne & Beaudichon, 1996).

1.3. The developmental transformation of private speech

Vygotsky’s (1934/1987) hypothesis regarding the progressive internalization of private speech has received considerable support (Winsler, Carlton, & Barry, 2000; Winsler, de León, Wallace, Carlton, & Willson-Quayle, 2003; Winsler & Naglieri, 2003). Its development follows an inverted U-shaped relationship with age, increasing and peaking in the preschool years and decreasing in the school years as it is progressively replaced by internalized speech. Winsler and Naglieri (2003) showed that private speech is most prevalent at age 5, but that silent, inner speech (inferred from self-reports) increases between ages 5 and 9, while private speech simultaneously diminishes. Based on the subtypes of private speech identified by Berk (1986), it has been shown that private speech becomes more task-relevant between ages 3.5 and 4.5 (Winsler et al., 2003).

Research on the internalization of private speech in the context of referential communication is sparse. Studies by Varenne and Beaudichon (1996) and Boada and Forns (2004) neglect the distinctions among private speech subtypes and between private and inner speech, making it difficult to analyze the pattern of internalization. Girbau’s (1997) study directly addressed this question and is the only referential communication study to develop additional indicators of inner speech. Based on the classical indicators of inner speech – i.e., whispering, inaudible muttering, and silent verbal lip movements (Berk, 1986; Winsler et al., 2005) – Girbau proposed that inner speech activity can be reliably inferred from silence if – and only if – the silence: (a) persists for 2 s or more; (b) is accompanied by facial behavior indicating concentration on the task; and (c) is not accompanied by communicative gestures directed to the interlocutor. She studied the emergence of private, inner, and social speech in 30 pairs of boys and girls and found no developmental decrease in private speech and its subtypes between 8 and 10 years of age from either speakers or listeners. Her results coincide with the pattern of private speech internalization in its later stages but do not allow us to draw conclusions about the preschool years.

1.4. Private speech and the cognitive demands on speakers and listeners in a referential communication task

A listener’s task is to comprehend the speech of others and a speaker’s task is to produce speech that is adapted for others and is unambiguous. In a referential communication task, both roles are critical for establishing effective communication (Dickson, 1981; Feigenbaum, 2009). However, the prospect for collective success is in greater jeopardy if the person occupying the speaker role fails to produce a well-formed and comprehensible message (Robinson, 1981). This factor ought to have consequences
for the quantity of private speech elicited according to communicative role. In one of the few referential communication studies to examine this issue, Girbau (1997) found that the production of private and inner speech varied in accordance with the speaker and listener roles: audible, task-relevant private speech utterances were significantly more frequent among speakers than among listeners, but only at age 8. Also, listeners produced more inner speech than speakers at ages 8 and 10. These findings suggest a possible difference in the timing of the internalization of private speech depending on (speaker vs. listener) role. However, Girbau did not analyze these differences as a function of the difficulty of the referent of communication.

1.5. Cognitive functions of private speech and relations to task difficulty

A diverse array of studies that used puzzles and academic tasks rather than communication tasks shows that private speech regularly increases in accordance with task difficulty – so long as the task is neither too difficult nor too easy, but instead is within the child’s zone of competence (Behrend, Rosengren, & Perlmutter, 1989; Fernyhough & Fradley, 2005; Winsler et al., 2000). In addition, these studies establish that as an individual masters a task there is a marked decrease in private speech (Duncan & Cheyne, 2002; Duncan & Pratt, 1997; Montero & de Dios, 2006). Varenne and Beaudichon (1996) and Boada and Forns (2004) extended the concept of task difficulty to referential communication tasks and were therefore able to provide empirical support for the cognitive role of private speech in the regulation of a communicative task, among both preschoolers and school-age children.

Boada and Forns (2004) focused their longitudinal study on the presence of private speech during speakers’ message-planning in a spatial referential communication task. To utter an unambiguous message about an object, a speaker has to identify it by name and, in some cases, by one of its attributes; he or she also must describe the location by means of deictic terms that express the spatial relations of support and proximity (Plumert, Ewert, & Spear, 1995). Support relations are concrete and refer to the particular location of a referent with respect to its physical contact with other objects (for example, “the ball on the floor”), whereas proximity relations are more abstract and refer to a general location of a referent without regard to any direct physical contact with other objects (for example, “on the left”). Boada and Forns found that 4-year-olds tend to formulate messages that only identify the referent and the support relation, while 6-year-olds construct messages that incorporate the abstract relations of proximity. Because structures of support are easier to understand than those of proximity (Bowerman, 1989; Plumert et al., 1995) it would be mostly around these latter relations that one would expect to observe private speech. Based on this reasoning, Boada and Forns defined private speech as only those speech fragments that occur in the middle of a communicative message and that express the more cognitively difficult information (proximity relations); if a fragment only expresses the cognitively easier information (support relations), however, it was not considered private speech. The private speech fragments that were thereby identified increased in amount between ages 4 and 6, and also increased in relation to the cognitive complexity of the referents at age 6.

1.6. The present study

This study has four basic aims: (1) to document the presence of private and inner speech in a challenging referential communication task using a new methodology for identifying private speech and inner speech; (2) to track the longitudinal development of private speech and inner speech using a referential communication task; (3) to explore the cognitive functions and development of private and inner speech in relation to the speaker and listener roles; and (4) to assess the cognitive functions of private and inner speech in relation to the level of cognitive difficulty of the task materials.

A major motivation for the work was to assess the practical utility of a rigorous new methodology that organizes all of the different classical and non-classical indicators of social, private, and inner speech into a standardized procedure for systematically applying these indicators to speech data in referential communicative contexts. A second motivation was to try to replicate the results of previous research using this classification system, as well as obtain new data that would raise fresh questions. Doing so has the potential to provide a fuller picture of the developmental progression and internalization of private speech in a referential communication task. Motivated by earlier findings
demonstrating some differences between the speaker and listener roles, our third aim is exploratory in character. We wondered if the quantity of private speech, its developmental progression, and its variation as a function of task difficulty are more closely related to one particular communicative role than to the other. Finally, the intention behind the fourth aim was to shed more light on the relation between task difficulty and private and inner speech. Would private speech increase with increasing difficulty of the referent and diminish when the referents are less challenging?

2. Method

2.1. Participants

The sample consisted of 20 Catalan-speaking children (10 girls) residing in Barcelona, Spain and attending a private, subsidized kindergarten. Ages ranged from 4-3 to 4-9 ($M=4.5$ years; $SD=2.26$). All attended school regularly. The socioeconomic level of the children's families – as defined by the parents' professions and academic levels – was considered upper middle-class. According to information provided by the school, 50% of the parents completed university level studies, 36% worked in administration and 21% in specialized technical fields. According to the teachers' assessments – based on their observations in the classroom and their knowledge of the children's academic history – the selected children had cognitive and linguistic levels within the normal ranges; these assessments were confirmed by the McCarthy test (1972). The cognitive level of the children at 4.5 years old was assessed individually one month before carrying out the communication test. The average test score on the general cognitive index (GCI) was 109.

Participants were drawn from three classrooms. Mixed-gender pairs were constructed of children drawn from the same classroom and who, according to the teachers' assessment, participated to a similar degree in speech activities. In each pair, one child was randomly assigned the role of speaker and the other of listener.

Participants were first assessed at the mean age of 4.5 years and retested at 6.5 and 8.5 years. The same pairs were maintained across assessments. Of original 15 pairs, only the 10 pairs available at all 3 sessions were included in the sample. No differences were detected between children included in the study and those lost through attrition.

2.2. Task

We used the Organization of a room task designed by Boada and Forns (1997). The task materials consist of two pictures and a set of eight moveable objects (referents). The speaker's picture shows a table, a bookcase and eight objects occupying fixed positions: two cups (one is big and red, the other is small and green), two bottles (one pink, and the other one green), a hat, a ball and two cats (one big, and the other small). Some objects are on the floor: the two cats, one beside the other, to the right of the table; the pink bottle to the left of the table and the ball underneath the table. Other objects are on top of the table: the hat to the right and the green bottle to the left. Finally, the two cups are on the bookcase in different spatial areas: the big one is above and to the right and the small one is below and to the left. The listener's picture contains only the drawing of the table and the bookcase, accompanied by a set of eight moveable objects identical to those in the speaker's picture, which the listener can place on the picture.

The task is designed to promote the use of spatial deictic terms that differ in degree of cognitive complexity based on the concrete or abstract qualities of the relations to which they refer – i.e., support or proximity. Level of difficulty is defined by the second of two basic acts comprising every verbal message: (a) the unambiguous identification of the referent by means of its name and, in some cases, one of its attributes; and (b) the unambiguous identification of the location of the referent by means of the concrete relations of support (top, middle, down) and the abstract relations of proximity (right, left) required to correctly identify each referent from among similar-looking objects located elsewhere in the picture. On the basis of these spatial relations, pictures can be designed in which particular referents are assigned a high or low level of cognitive difficulty (Boada & Forns, 2004). The difficult referents were those that required greater use of proximity relations in order to locate them.
correctly. For example, the placement of the big cat involves communicating that it is on the floor (support relation) to the right of the table (one proximity relation) next to the small cat (a second proximity relation), whereas the placement of the hat only involves conveying that it is on top of the table (support relation) to the right (one proximity relation). Referents of high cognitive difficulty were the cups, the cats and the pink bottle, and the referents designated as cognitively easier were the hat, the ball and the green bottle.

2.3. Procedure

The procedure was based on the classic method of Krauss and Glucksberg (1969) with certain changes made to enhance ecological validity (Lloyd, Boada, & Forns, 1992). Pairs performed the task in rooms at their schools separated from their usual classrooms. First, the adult randomly assigned the roles of speaker and listener and presented the task as a game. Then, an opaque screen was placed between the children so that they could not make direct eye contact. The speaker's objective was to describe – using vocal communication only – the arrangement of objects pictured in the room, while the listener's objective was to recreate that arrangement in his/her own picture of the room by placing the moveable objects appropriately. Their common goal was to produce two identical pictures. The adult's role was to maintain the communicative channel between the two children, and to re-establish it in case of interruption, offering any verbal support considered necessary to complete the task.

The task had no time limit. It was considered completed when all of the objects had been named and positioned. The task was repeated three more times, for a total of four trials per assessment period. At the end of each trial, the adult showed the speaker how the partner had positioned the objects. Both children were then immediately invited to try again to improve their performance.

This procedure was identical at the other two assessment periods, two and four years later.

2.4. Data analysis

Classification system: The sessions were recorded on audiotape and videotape and then transcribed and coded. The speech was then classified according to the system presented in Table 1.

Presence or absence of private speech was assessed in both the messages by the speaker and the responding actions by the listener. Following Boada and Forns (2004), messages were defined as utterances whose content was related to identifying an object and its position, and actions were defined as the corresponding movements of picking up an object and placing it on the picture in response to the speaker's message. From the messages produced within an extended conversational sequence related to the same object, only the last message (and corresponding action) in a conversational sequence was analyzed for the presence (or absence) of private and inner speech. The last turn was the one that contained the most complete information because it embodied the accumulated knowledge and experience gained from the previous turns, and because it was the turn that demonstrated the most agreement between the interlocutors. For example, in the sequence: Speaker: “The cat is on the floor.” Listener: “The small one?” Speaker: “Yes, the small one on the floor”, the presence of private speech was considered only in relation to the speaker's last turn at talk.

To counterbalance the effects of practice, familiarity, and fatigue across trials, data were collected from the first and third trials only; these data were combined. The first trial was chosen over the second because practice and familiarity with the task were at a minimum, and the third trial was chosen over the fourth because the children were less fatigued.

Criteria for distinguishing private speech and its subtypes from social speech: The speakers’ messages and the listeners’ actions were assessed separately for occurrence of private speech. The subtypes of private speech were identified according to how audible the utterances were, and how well they related to, and focused on, the task (Berk, 1986). Because fully developed inner speech cannot be directly observed – that is, only the transitional forms leading up to it have observable properties such as muttering and lip movements – here we classified inner speech as a “silent or inaudible” subtype of private speech. Thus, we divided private speech into two basic subtypes: audible (task relevant or distracted) and silent private speech, the latter consisting mainly of the external manifestations of inner speech (see Table 1).
Table 1
Classification of subtypes of private speech used in a referential communication task.

<table>
<thead>
<tr>
<th>Subtypes of private speech</th>
<th>Examples of classification categories</th>
<th>Speaker</th>
<th>Listener</th>
</tr>
</thead>
</table>
| Audible private speech     |  **Messages with task-relevant private speech**  
Messages that include the repetition of some verbal elements or that contain verbal elements that complete or repair previous information in the same message.  
The repetitions and repairs are not immediate, but delayed: accompanied by pauses and pause-fillers. This includes verbalizations with explicit semantic content about the task of formulating the message  
**Actions with task-relevant private speech**  
Actions that include verbalizations that indicate doubts about the ongoing action: exclamations and/or pause-fillers. Also, actions that contain verbal elements that indicate a repair of the action, or that focus semantically on some aspect of the ongoing action  
*For both messages and actions*, these verbalizations are made in the absence of eye contact with the adult, or communicative gestures, and are made simultaneously with one of the following indicators: (a) a low tone of voice (b) nonverbal behavior that indicates concentration on the task (looks, gestures, etc.). | “the bottle, oh no!” [said without looking at the adult, but while looking with concentration at the picture]  
“The bottle at (pause) at (pause) at the side of the table” [this is repeated while looking with concentration at the picture and pointing with a finger at the bottle] | × |  |
| Distracted                 |  **Messages with distracted private speech**  
Messages with verbal expressions that indicate a loss of concentration or attention to the task | “The hat-oops!” [says “oops” when another object fails on the floor] | × |  |
<table>
<thead>
<tr>
<th>Subtypes of private speech</th>
<th>Examples of classification categories</th>
<th>Speaker</th>
<th>Listener</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actions with distracted private speech</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actions with verbal expressions that indicate a loss of concentration or attention in the process of positioning the referent.</td>
<td>&quot;oh. oh&quot; [playing with the objects while the speaker formulates the message]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>For both messages and actions</strong>, these verbalizations are made in the absence of eye contact with the adult, or communicative gestures, and are made in a low tone of voice. In addition, nonverbal aspects are considered when identifying a loss of concentration (direction and focus of gaze and body movements).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Messages with silence (inner speech)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Messages with silence: messages with pauses longer than or equal to 2 s that are unaccompanied by communicative gestures. Excluded are silences occurring during contact with the interlocutor or immediately following the interlocutor’s utterance</td>
<td>&quot;The pink bottle at the side of the table on the, on the [pause &gt; 2 s left] during the silence the child looks with concentration at the picture and signals to the left, and afterwards verbalizes that the bottle is to the left]&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Actions with silence (inner speech)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actions with pauses longer than or equal to 2 s where a child simultaneously shows, through his/her gestures, a clear modification or repair in the action aimed at positioning the referent</td>
<td>&quot;The pink bottle on the side&quot; (After this message): Listener: [silence] [during this silence: looks intently at the picture, takes the pink bottle and puts it to one side of the table, looks toward the other side, looks up, and then repositions the bottle on the other side]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>For both messages and actions</strong>, these silences occur simultaneously with the following indicators: (a) nonverbal conduct that indicates concentration on the task; (b) absence of eye contact with the adult or communicative gestures.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Private speech in speakers’ messages: To identify private speech occurring within speakers’ social speech messages, we integrated the proposals made by Boada and Forns (2004), Girbau (1997), Frawley and Lantuof (1985) and Varene and Beaudichon (1996). A message was considered to include private speech if it contained certain types of audible verbal elements and/or silences that were not explicitly directed at another person and that indicated self-regulation in the task of formulating the message. To ensure the intrapersonal function of these elements, it was essential to consider linguistic, paralinguistic and nonlinguistic criteria, as described below.

Audible private speech was identified by audible verbal elements that included certain dysfluencies that disrupted the flow of the message. Specifically, these dysfluencies consisted of repetitions and corrections that were always accompanied by pauses and pause-fillers (Manfra, Tyler, Shiflett & Winsler, 2003). Immediate repetitions or corrections were excluded (e.g. “the r-cat”), as were grammatical or pronunciation errors. In addition, verbal elements with explicit semantic content about the child’s attempts to gain control over the task of formulating the message were part of the criteria (e.g., “the cup in . . . [pause] . . . what's it called? [the question is uttered softly] . . . in the bookcase”). All of these verbal features were considered indicators of private speech only if their intrapersonal nature could also be justified by evidence, such as the absence of eye contact with the adult when these utterances were made, and by occurring at the same time as at least one of the following indicators: (a) a low tone of voice; and/or (b) nonverbal behavior that indicated an explicit concentration on the task (e.g., the child closed his/her eyes while formulating the message, or covered them with his/her hands). With regard to the function of private speech in relation to the task, we distinguished between task-relevant and distracted private speech.

With respect to the external manifestations of inner speech, or “silent” private speech, we integrated the proposals made by Berk (1986) and Girbau (1997) and inferred inner speech activity only when silences lasting 2 s or longer were embedded within a message and were accompanied at some point during the pause by inaudible muttering or silent lip movements. In addition, there could be no simultaneous communicative gestures or eye contact with the interlocutor, and one of the aforementioned aspects of nonverbal behavior, indicating concentration on the task, also had to be observed simultaneously.

Private speech in listeners’ actions: Private speech utterances produced by listeners when they were complying with speakers’ messages were classified on the basis of the standard criteria for identifying private speech in a manipulative sequence (Winsler et al., 2005). Following these criteria, an action was considered to include private speech if it contained certain audible verbal elements and/or silences that did not appear to be explicitly aimed at others present, and that indicated self-regulation in the action of positioning the referent object on the picture. These elements occurred at the same time as the action or sometimes immediately before or after it. As in the case of the speaker, linguistic, paralinguistic, and nonlinguistic criteria had to be considered in order to identify private and inner speech.

Audible verbal elements produced by listeners were coded as audible private speech if they included certain dysfluencies that indicated doubts or hesitation with respect to the action; thus, the action was not carried out immediately after listening to the message (e.g., after receiving the message the listener said: “Yes, [pause] um”. While she said “yes” she picked up the referent object, and when she said “um” she explored the picture with concentration and finally positioned the object). Similarly, dysfluencies that indicated corrective action were considered private speech (e.g., the listener said: “OK, oh no! There.” He said “oh no” in a low voice while he removed the object from its original position to put it in another place). In addition, we also considered as private speech any utterances that were focused on some aspects of the task and that had a semantic content that was relevant to placing the object (e.g., the child said “here” while placing the object). A verbal element was identified as private speech if there was no eye contact with the adult and if it occurred at the same time as at least one of the following indicators: (a) low tone of voice, and (b) nonverbal behavior that indicated concentration on the task or difficulties in placing the object. As before, we distinguished between task-relevant and distracted private speech.

With regard to “silent” private speech produced by listeners as they complied with speakers’ messages, we considered as possible candidates only those silences or pauses that persisted for 2 s or longer. Furthermore, for these within-performance silences to be considered as indicators of inner
speech, inaudible muttering or silent lip movements had to be present at some point during the silence, and there could be no communicative gestures or eye contact with the adult. In addition, one of the aforementioned aspects of nonverbal behavior had to occur, indicating concentration on the task or a difficulty in attempting to master the task, and the child had to simultaneously show, through his/her gestures, a clear modification or repair in the action aimed at positioning the referent.

Coding reliability: Two independent coders first underwent a training session in which they were provided the categorization system with the operational definitions and examples, three coded transcriptions for three pairs of children, and the respective videotapes. They were then asked to code another three transcripts (10% of the total sample) selected randomly and accompanied by the corresponding videotapes. For speakers’ messages, agreement over the presence or absence of private speech utterances was 84.44% ($k = 0.69$); for listeners’ actions, agreement was 86.67% ($k = 0.73$). Regarding the subtypes of private speech and inner speech, agreement over classifying messages was 81.25% ($k = 0.72$), while for actions the agreement was 80% ($k = 0.70$).

### 3. Results

A total of 156 messages were identified at 4.5 years, 160 messages at 6.5 years and 158 at 8.5 years. The difference across age was not significant. Nor were the total of 135 actions by listeners identified at 4.5 years, 138 actions at age 6.5 and 142 at age 8.5.

Frequencies of messages containing private speech and inner speech by age are shown in Table 2. There was a low occurrence of distracted private speech (always below 1% of the total) in the messages and actions at all ages. Therefore, distracted private speech was excluded from all subsequent analyses, in effect making the category audible private speech synonymous with the category task-relevant private speech.

#### 3.1. Messages

At age 4.5, most messages did not contain private speech (64.74% of the 156 messages). When it did appear in this age group, it almost exclusively took the form of audible private speech, with inner speech being practically absent. At 6.5 years, however, most messages did contain private speech (76.25% of 160 messages). Again, audible private speech predominated (Table 2). There occurred significantly more messages with audible private speech than inner speech (Wilcoxon test, $z = -2.818$, $p = .005$). Number of messages containing private speech had increased significantly from ages 4.5 to 6.5 ($U = 6.500, p < .001$), as had those with audible private speech ($U = 7.500, p < .001$) and those with silent, inner speech ($U = 12.500, p = .002$).

At age 8.5, messages containing private speech represented 83.54% of 158 messages. The most common category was that of messages containing audible private speech (Table 2). Messages with audible private speech occurred more frequently than messages with inner speech (Wilcoxon test, $z = -2.814$, $p = .005$; $z = -2.661$, $p = .008$). Number of messages containing private speech or audible private speech did not change significantly from 6.5 to 8.5 years. However, the increase in the mean number of messages with inner speech approached significance ($U = 26.000, p = .065$).

#### 3.2. Actions

At 4.5 years, only a small percentage of listeners’ actions contained any private speech (Table 3). By ages 6.5 and 8.5, this percentage had increased to almost 30%. This increase was significant between ages 4.5 and 6.5 ($U = 10.000, p = .002$). At age 6.5 most actions were carried out without private speech. When it did appear, the most common form was audible private speech, followed by inner speech. There were significantly more actions with audible private speech than with inner speech (Wilcoxon test, $z = -2.692$, $p = .007$). Compared to age 4.5, there was significantly more audible private speech ($U = 11.500, p = .002$) and inner speech ($U = 20.000, p = .005$) in listeners’ actions. At 8.5 years old, actions continued to be carried out mainly without private speech. When it did appear, the most common category continued to be that of audible private speech. There were no significant differences with respect to the production of audible private speech and inner speech. The mean number of instances of
**Table 2**
Number of messages containing private speech and inner speech by age.

<table>
<thead>
<tr>
<th>Age</th>
<th>Messages with private speech</th>
<th>Messages without private speech</th>
<th>Total messages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Task-relevant private speech</td>
<td>Silent (inner speech)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td></td>
</tr>
<tr>
<td>4.5 years</td>
<td>5.30 (3.27)</td>
<td>0.10 (0.32)</td>
<td>5.50 (3.44)</td>
</tr>
<tr>
<td>n = 10</td>
<td>33.97</td>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>6.5 years</td>
<td>10.70 (2.00)</td>
<td>1.50 (1.18)</td>
<td>12.20 (2.62)</td>
</tr>
<tr>
<td>n = 10</td>
<td>66.87</td>
<td>9.37</td>
<td>0.00</td>
</tr>
<tr>
<td>8.5 years</td>
<td>10.50 (2.71)</td>
<td>2.70 (1.63)</td>
<td>13.20 (3.12)</td>
</tr>
<tr>
<td>n = 10</td>
<td>66.45</td>
<td>17.08</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
</tbody>
</table>
inner speech in listeners’ actions was higher than at 6.5 years. This increase was significant \(U = 23.000, p = .029\).

### 3.3. Difficulty of the referent

At all ages, there was a higher frequency of messages with private speech for the difficult referents than for the easy referents (Table 4), \(\chi^2(1) = 7.263, p = .007\) at 4.5 years; \(\chi^2(1) = 17.743, p < .001\) at 6.5 years; \(\chi^2(1) = 21.710, p < .001\) at 8.5 years. By age 8.5, almost all messages about difficult referents contained audible private or inner speech (compared to 66.67% for easy referents).

The relation between referent difficulty and audible private and inner speech in messages was significant at 6.5 and 8.5 years, \(\chi^2(1) = 1.208, p = .2717; \chi^2(1) = 0.002, p = .9664\), respectively. (This relation was not examined in the 4.5-year-old group, given the low frequencies.) No relations were significant for actions.

### 4. Discussion

Guided by Vygotskian theory, a major innovation of this study is establishment of a coding system for utterances of private speech and inner speech that occur in children’s referential communications with others. The application of this system in a longitudinal study allows us to identify a developmental sequence in an area of private speech investigation that is fairly new – the regulation of referential communication.

The corpus of private and inner speech that was produced at all ages in speakers’ messages and listeners’ actions was large enough that statistical significance tests could be applied to task-relevant private and inner speech at each age level, with the exception of listeners’ actions at 4.5 years of age. The finding of robust quantities of private speech replicates the results obtained in studies that employed the same communicative task. What is novel here is the system of categorization, which
is both broader and more stringent, that was able to detect a greater quantity of private and inner speech. Boada and Forns (2004) found that a quarter of the messages produced by their participants at age 6 contained audible private speech, compared to the two-thirds containing audible private speech that we detected in participants of the same age.

The pattern of development and internalization of private speech found in previous referential communication studies is based on a fairly narrow age span. Our longitudinal study broadens the age range to 4–8 years of age. The most notable increase in private speech occurred between 4.5 and 6.5 years of age for both speakers and listeners. This finding replicates the one obtained by Boada and Forns (2004), although they only considered the audible private speech produced by speakers. The amount of private speech at age 6.5 was substantial and remained stable between 6.5 and 8.5 years. At these ages, speakers produced a large amount of private speech (over three quarters of message statements at 6.5 and 8.5 years). Regarding developments between 6.5 and 8.5 years, existing studies do not provide a basis for comparison. Nevertheless, our findings are consistent with previous research on private speech in non–communicative tasks (Winsler et al., 2000, 2003; Winsler & Naglieri, 2003).

According to Vygotsky (1934/1987), the early increase and subsequent stabilization of private speech indicates the beginnings of autonomous regulation in performing a task, which in the present case involved both a communicative and a cognitive component (Wertsch, 1985). Our results are in accordance with studies that consider the early school years (around age 6) key to emergence of the ability to regulate communication (Boada & Forns, 1997, 2004; Lefebvre-Pinard, 1985).

We detected an increase in inner speech with age, both in speakers’ messages and listeners’ actions. This finding is of particular relevance given few studies make use of inaudible external concomitants of inner speech in communicative situations (Diaz, 1992; Girbau, 1997, 2002). We found that silences with inaudible muttering or silent lip movements are almost nonexistent at age 4.5, and increase with age, as other studies undertaken from a Vygotskian point of view have found using tasks other than those involving referential communication (Winsler & Naglieri, 2003). However, our results do not allow us to infer a perfect sequence of internalization, given the increase in inner speech between 6.5 and 8.5 was not statistically significant and was not accompanied by a decrease in task–relevant private speech. Our results could be viewed as the developmental antecedent to the outcomes obtained by Girbau (1997) in children 8–10 years of age.

Regarding the communicative role, we found shared and divergent trends at each age level. The developmental sequence for both interlocutors is similar. However, at every age, the proportion of messages with private speech is greater than the proportion of actions containing private speech. In addition, at age 8, the proportion of inner speech produced by listeners (of the number of actions containing private speech) was about twice as large as it was for speech produced by speakers (of the number of messages containing private speech). Finally, at every age, audible private speech predominated over inner speech, with the exception of listeners at age 8, who displayed no significant difference.

These findings are in line with those of Girbau (1997) and suggest that the development of inner speech may be slightly more advanced in listeners than it is in speakers. This interpretation fits with the view of Robinson (1981), who argued that referential communications tasks – by their very structure – place responsibility for producing a successful message on the person in the speaker role. In other words, the task could be more challenging for the speaker than for the listener. In any case, the data presented here suggest, but do not confirm, a difference in timing between speakers’ and listeners’ internalization of private speech. They also show that the emergence and internalization of private speech are complex and related to both age and task demands. Future researchers could take advantage of the methodology presented in this study to explore these differences in patterns of development in speaking and listening skills.

With respect to task difficulty, results revealed more private speech in messages regarding difficult referents than easy ones, in accordance with the results expected from a Vygotskian perspective. At age 4.5, the cognitive function of private speech in message formulation is only foreshadowed: although difficult referents generate more private speech than easy ones, most messages do not contain private speech. However, at 8.5 years old, almost all messages about difficult referents contain private speech. Although we did not find a relation between private speech in the listeners’ actions and the level of difficulty of the referent, we wonder if, for the listener, the cognitive challenge is not so much about
the referent per se, but about the adequacy and degree of ambiguity of the message being received. This is an issue that needs further exploration.

These results corroborate and extend those obtained by Boada and Forns (2004) regarding the increase in private speech at 6 years and those of Varenne and Beaudichon (1996) regarding the continuation of the relation to task difficulty to age 8. Our results also parallel those found in studies of private speech in non-communication tasks, in which a relation between task difficulty and the emergence of private speech was reported (Behrend et al., 1989; Fernyhough & Fradley, 2005; Winsler et al., 2000). However, we were not able to establish a curvilinear relation between task difficulty and private speech, as we only worked with two levels of difficulty of the referents.

Studies have shown that less private speech is produced as an individual masters a task (Duncan & Cheyne, 2002; Duncan & Pratt, 1997; Montero & de Dios, 2006). When describing easy referents, a reduction in private speech could be expected with increasing age as the referents are progressively mastered. Therefore, our discovery of an increase with age in amount and types of private speech related to both difficult and easy referents poses a theoretical challenge, obliging us to further consider the sources of cognitive difficulty posed by the referential communication task.

Referential communication poses at least two kinds of cognitive challenge for children. One is to learn how to put into words the relations (or spatial concepts) of support and proximity for one’s own understanding. The other is to learn how to communicate this understanding to others. To unambiguously communicate the location of easy and difficult referents, it is necessary to understand the meaning of the relations of support and proximity. Support relations are incorporated into messages at age 4, while proximity relations are cognitively more complex and are only partially incorporated into messages at ages 6 and 8 (Boada & Forns, 2004; Bowerman, 1989; Plumert et al., 1995). Private speech is crucially important to this effort because it assists children to understand these concepts. At 6 years old, the attempt to describe relations of proximity seems to represent a new source of difficulty as much for easy referents as for difficult ones, with the level of difficulty even greater in the latter case. Thus private speech increases with age for both difficulty levels.

In a referential communication paradigm messages are generally analyzed in terms of degree of adequacy necessary for the listener to solve the task (Dickson, 1981). By identifying private speech in the messages, the analysis can be expanded to include elements that can be interpreted as an expression of a child’s intention to regulate the communication (San Martín et al., 2009). For example, children’s private speech typically omits mention of the topic of conversation (because it is already known), but in order to communicate their understanding to others, children must become conscious of the need to include additional information or their speech will remain incomprehensible to their audience. Hence the age increase in private speech may reflect the continuing cognitive challenges associated with both learning to think and learning to communicate unambiguously about spatial deictic terms. This is consistent with studies showing that the mastery of communication is progressive but lags behind other cognitive achievements (Camaioni et al., 1998; Lloyd et al., 1998). Thus, the study of private speech can be a valuable tool in enhancing our understanding of how children respond cognitively to the social demands of a communication task (Varenne & Beaudichon, 1996; Wertsch, 1985).

Feigenbaum (2002, 2009) has urged that future research explore the role of private and inner speech in the development of discourse skills. Of interest would be the consequences that follow from manipulating the experimental conditions within the referential communication task. Specifically, the kind of feedback that speakers receive from listeners could be varied systematically to determine its effect on speakers’ private and social speech in subsequent iterations of the task.

In conclusion, this longitudinal study provides evidence to support several psycholinguistic claims: that private and inner speech are intimately involved in referential communicative situations; that they may serve different communicative and cognitive purposes at different ages; and that they vary in amount and function according to communicative role and increasing complexity of the material to be communicated. In addition, the present research suggests that dysfluent speech occurring within discourse is very likely a manifestation of verbal thinking that takes the form of private or inner speech. But perhaps the most novel and useful contribution of this study is the suggestion that private and inner speech embedded in discourse may serve a regulatory role in social speech communication.


