Spanish-speaking children do not always overuse *estar*

Los niños hispanohablantes no siempre hacen uso excesivo de 'estar'

Carolina Holtheuer
choltheuer@ciae.uchile.cl
Universidad de Chile
Chile

Abstract: Current research has shown that young Spanish-speaking children distinguish between *ser* and *estar* but fail to provide clear answers to questions such as which copula is easier/first to acquire. Some studies show that children perform worse with *estar* while others show that *ser* is mastered with more difficulty than *estar*. The current study presents the results of two experiments that tested children's comprehension of *ser* and *estar*. The results indicate that i) children compute the inferences associated with *estar* when natural contextual information is provided, ii) children's copular distinction is supported by syntactic factors, iii) children perform poorly when copula choice depends on contextual factors alone, and most importantly, iv) children do not assign *estar* interpretations to *ser* compatible environments. It is concluded that Spanish-speaking children do not always overuse *estar* as opposed to *ser*.

Resumen: Investigaciones recientes han mostrado que los niños hispanohablantes distinguen entre 'ser' y 'estar' pero no han respondido con claridad a preguntas básicas como cuál cópula se adquiere primero o con mayor facilidad. Mientras algunos estudios indican que los niños tienen más problemas con la adquisición de 'estar', otros muestran que 'ser' es más difícil de dominar. En este estudio se presentan los resultados de dos experimentos que investigan la comprensión de 'ser' y 'estar'. Los resultados obtenidos indican que: i) los niños procesan las inferencias asociadas a 'estar' cuando se les presenta información contextual de manera natural, ii) la distinción entre 'ser' y 'estar' se apoya en factores sintácticos, iii) los niños tienen mayores problemas cuando la elección entre 'ser' y 'estar' depende solamente de factores contextuales, y (iv) los niños no asignan interpretaciones con 'estar' a situaciones en las que el uso de 'ser' es más apropiado. Se concluye que los niños hispanohablantes no siempre abusan de la forma 'estar' en comparación con 'ser'.

Palabras Clave: 'Ser' y 'estar', adquisición de lengua materna, tareas de comprensión.

INTRODUCTION

The Spanish copulas *ser* and *estar* have received great attention from theoretical linguists and second language acquisition experts but relatively little is known about how they are acquired by Spanish speakers. With the exception of Silva-Corvalán and Montanari's (2008) study of *ser* and *estar* acquisition by an English-Spanish bilingual child, in which the acquisition of *estar* is reported to be delayed, it is generally accepted that Spanish-speaking children master the basic properties of the copulas and use them in an adult-like fashion. However, studies that test comprehension of the copulas with adjectives have shown that 4- to 6-year-old children (Schmitt, Holtheuer & Miller, 2004; Schmitt & Miller, 2007) have problems with the pragmatic dimension of the copulas. In addition, a recent study of natural production has revealed that 2- to 3-year-old children make some morpho-syntactic errors while using *ser* and *estar* (Holtheuer, 2009). Hence, acknowledging that children are highly accurate users of *ser* and *estar* should not lead us to think that children's early production and comprehension of the copulas are effortless and adult-like.

Up to date research indicates that young children make a general distinction between *ser* and *estar* but fail to provide clear answers to basic questions such as which copula is easier/first to acquire and whether children default to one copula in their early productions. While some studies show that children perform worse with *estar* than with *ser* (Sera, 1992; Schmitt et al., 2004) other show that *ser* is mastered with more difficulty than *estar* (Sera, Bales & del Castillo, 1997; Schmitt & Miller, 2007).

In order to expand current empirical findings this study presents the results of two experiments, a Picture Selection Task (hence PST) and an Acceptability Task (hence AT) that tested children comprehension of *ser* and *estar* sentences in controlled discursive contexts. The objectives of the PST are (i) to examine whether Spanish-speaking children know that *estar* as opposed to *ser* is used to describe contextually marked properties and (ii) to investigate whether children learn *ser* and *estar* by keeping track of the type of adjective (*ser* adjectives or *estar* adjective). The objectives of the AT are (i) to investigate what kinds of information (syntactic, semantic and pragmatic) children use for distinguishing *ser* and *estar* and (ii) to examine whether children experience more difficulty with one or the other copula. The results indicate that Spanish-speaking children know the basic properties of the copulas but are still developing their acquisition of the pragmatic properties of *ser* and *estar*. Crucially, both experiments indicate that children do not assign *estar* interpretations to *ser* compatible environments and therefore do not support the idea that children overuse *estar*.

1. Distribution and basic properties of *ser* and *estar*

Ser and *estar* are present in different syntactic constructions (e.g. both *ser* and *estar* take adjectival and prepositional phrases as complements, only nominal phrases appear with *ser* but not *estar*, and only *estar* is the auxiliary for the progressive form). However, it is the 'copula + adjective' construction the one that has attracted most interest from researchers because it is in this domain where choosing the appropriate copula becomes a complex task that involves knowledge of syntactic, semantic, pragmatic and discursive factors. Notice that while

many adjectives are compatible with both *ser* and *estar*, not all adjectives can appear with both copulas (* stands for ungrammatical and? for infelicitous sentences):

(1)
a. Mariana es/está linda/callada. [M. is pretty/ quiet]
b. Carla*es/está embarazada/cansada. [C. is pregnant/tired]
c. Maite es/? está inteligente/sincera. [M. is intelligent/sincere]
d. Juan es/está vivo. [J. is intelligent/ alive]
e. Juan es/está alto. [J. is normally tall/ tall right now]

Moreover, while some adjectives have different lexical meanings depending on whether they appear with *ser* or *estar* (1d), other adjectives keep the same lexical meaning but yield a different interpretation of the predicate (1e).

Although many adjectives are compatible with both copulas, their distribution is dictated by a complex combination of syntactic, semantic and discursive-pragmatic factors that make it difficult to predict and explain their use. This difficulty is embodied by the numerous proposals that have attempted to explain the copula distribution in Spanish (Bello, 1951; Gili Gaya, 1955; Luján, 1981; Clements, 1988, 2005; Diesing, 1990, 1992; Schmitt, 1992, 1996, 2005; Maienborn, 2003, 2005; Arche, 2006).

Traditionally the difference between *ser* and *estar* has been casted in terms of permanent or inherent and temporary or non-inherent properties (Bello, 1951; Gili Gaya, 1955; Vañó-Cerdá, 1982), but it is now widely accepted by most linguists that temporal notions alone cannot determine the *ser* and *estar* distribution (Schmitt, 1992, 1996, 2005; Maienborn, 2003, 2005; Arche, 2006). Here, we follow Schmitt's (2005) proposal that views *ser* and *estar* as the exponents of the permanent and temporary copula as the result of a pragmatic phenomenon. The two copulas are semantically distinct: while *ser* is underspecified for aspect, *estar* is specified for aspectual properties. Since *ser* is aspectually unspecified it gives rise to an implication that a property holds independent of time while *estar* is a state that holds at a certain time (time t), and therefore gives rise to an implicature of temporariness.

2. Child language acquisition of *ser* and *estar*. Previous research

Sera (1992) is the first comprehensive study of the acquisition of the Spanish copulas that shows that children master the distribution of *ser* and *estar* with different predicates. Sera (1992) examined a Spontaneous Speech sample of two boys (ages 1;6 and 3;6) and an Elicited Production sample from eleven 3-year-olds, twelve 4- year-olds, eleven 5-year-olds, twelve 9-year-olds and five adults. In her study, children used *estar* with locative predicates and *ser* with nominal predicates in an adult manner. However, the children differed from adults in two important ways: first, they overused *estar* as compared with *ser* in adjectival predicates and second, they incorrectly used *estar* in locative predicates with event subjects (*estar* is used to locate objects and *ser* to locate events). Given that children did not distinguish between the two copulas based on the semantic properties of the subject (object or event), Sera (1992) suggested that children's *ser* and *estar* choice may be determined by syntactic factors.

In another study, Sera et al. (1997) tested whether the use of *ser* and *estar* affected the ability of Spanish-speaking children (age range 2;11- 5;10) to identify real and apparent properties. Even though several children used *ser* for describing temporary properties, they concluded that Spanish-speaking children correctly matched *ser* with permanent properties and *estar* with apparent properties. Unfortunately, in this last study no adults were included and therefore we cannot compare adults' and children's performance.

More recently, Schmitt et al. (2004) examined whether children allow pragmatically appropriate interpretations to *ser* and *estar* by using a Picture Matching Task (PMT) and an Acceptability Judgment Task (AJT). There were twenty children (mean age 4;6) in the PMT and sixteen children (mean age 4;4) in the AJT. In the PMT, the child had to choose between two pictures, one depicting a canonical (permanent) property of a character such as a tall giraffe and the other depicting a non-canonical (temporary) property such as a short giraffe standing on top of a table. The participants were asked to point to one of the pictures as an answer to the experimental question (e.g. Which is (*ser/estar*) tall?). It was found that although children distinguished between the two copulas, unlike adults, they assigned *ser*

interpretations to questions with *estar*. The authors concluded that children's problems may have been due to their pragmatic inability to rule out *estar* in typical *ser* contexts. In their other experiment, the AJT, participants judged whether the *ser/estar* experimental sentence was acceptable in three different conditions that controlled for the kind of information provided: syntactic, semantic and discursive-pragmatic. The children performed unlike adults in all three conditions but performed worst in the condition that controlled for discursive-pragmatic information. Like in the PMT, the children were more willing to accept *estar* in *ser* appropriate contexts than vice-versa suggesting that children overuse *estar*.

In another study, Schmitt and Miller (2007) showed that Chilean and Mexican Spanish speaking children understand some pragmatic properties associated with *ser* and *estar*. In one task involving only color adjectives, children (4;5–6;3, mean age: 5;3) approached adult levels in the use of both *ser* and *estar* but unlike adults, and similarly to Schmitt et al. (2004), children still allowed *estar* to be associated with permanent properties. In another task with scalar adjectives (e.g. 'tall'), children (4;7–6;0, mean age: 5;5) patterned with adults in the *estar* but not in the *ser* condition.

The research findings so far present us with complex and inconsistent data with regard to which copula is overused and/or acquired first. Even though the above studies are not easily comparable (different methods and different tasks), it is obvious that we need to carry out more experiments to understand better how *ser* and *estar* are acquired by monolingual speakers. The present study adds to the current body of research by presenting two novel experiments that tackle the comprehension of *ser* and *estar* with adjectives in Chilean and Cuban Spanish.

3. Studies

In order to further explore children's understanding of the copulas we carried out two experiments. The first tested comprehension of *estar* by using a Picture Selection Task (PST) design and the second experiment tested the use of both *ser* and *estar* in specific contexts by using a version of McDaniel and Cairns' (1996) Grammaticality Judgment Task, an Acceptability Judgment Task (AT). Consistent with the Semantic Subset Principle (Crain & Thornton,

1998) it was hypothesized that children may acquire *estar* before *ser*. This principle leads children acquiring language to first choose the representation that is true in the smallest set of circumstances. Since *estar* can be a subset of *ser* (Luján, 1981), we predicted that if children follow the Semantic Subset Principle, they should acquire *estar* before *ser*. In addition, following general experimental findings that children perform significantly worse when the interpretation of linguistic terms requires knowledge of the interaction of syntactic & discourse-related constraints (Avrutin & Wexler, 1992; Noveck, 2001; Papafragou & Musolino, 2003; Schmitt et al., 2004), it was predicted that children would have more problems when the copula choice depends on discourse factors alone.

3.1. The Picture Selection Task: Comprehension of estar

3.1.1. Goal of the experiment

The goal of the PST was two-fold. First, we wanted to test whether Spanish speaking children know that *estar* is connected to the discourse and gives rise to marked interpretations of the subject in adjectival predicative constructions. It is known that *estar* differs from *ser* in that the former is always linked to the discourse while the latter is not (Maienborn, 2003, 2005; Clements, 2005; Schmitt, 2005; Arche, 2006). If children know the discursive-pragmatic properties of *ser* and *estar* they should consistently select the picture that is more appropriate to match a marked interpretation of the predicate. Alternatively, we would expect two outcomes: either children will select any picture at random (therefore, showing no preference for the specific *estar* representation) or they will choose the picture that is consistent with a *ser* interpretation.

Our second goal was to find out whether children use typicality cues. Sera (1992) reported that children use certain adjectives with *ser* and a different set of adjectives with *estar* and suggested that children might use adjective class (*ser* adjectives/ *estar* adjectives) in order to learn the distribution of the copulas. In order to control for typicality effects we used experimental sentences that only contained *estar* with adjectives that typically appear with *ser*. The adjectives used in this task were selected based on a norming study in which 19 adult native speakers of Spanish had to fill in the blanks of adjective predicates with either *ser* or *estar* (see

Appendix A for items used in the norming study). Since the experimental sentences included those adjectives that typically occurred with copula *ser* in the production of adults, by using *estar*, we controlled for the potential children's preference for allowing typical *ser* adjectives to appear only with *ser*.

3.1.2. Participants

Thirty Cuban-Spanish speaking children with mean age of 60.7 months, an age range of 39-87 months, SD = 11.8 (mean age 5;0, range 3;3-7;3) and eighteen Chilean-Spanish speaking children with a mean age of 58.5 months, an age range of 49-71 months, SD = 7.3 (mean age 4;10, range 4;1- 5;11) participated in this task. There were five 3-year-olds, six 4-year-olds, fourteen 5-year-olds, four 6-year-olds and one 7-year-old in the Cuban sample (14 boys and 16 girls). There were twelve 4-year-olds and six 5-year-olds in the Chilean sample (8 boys and 10 girls). The Cuban children attended public kindergartens located in Havana City and came from a mixed socioeconomic background. The Chilean sample consisted of middle class children that attended a private kindergarten center in Punta Arenas. In addition, twenty-five Cuban adults and nine Chilean adults were tested as controls. The difference in the number of subjects between the Cuban and Chilean groups is not a methodological problem since what we are testing is the difference in performance between children and adults that share the same dialect.

The reason for including two dialectal varieties of Spanish is that prior sociolinguistic research has shown that there is variation in the use of the copulas among Spanish dialects (Silva-Corvalán, 1991, 1994; Gutiérrez, 1992; De Jonge, 1993; Fernández, 1999). In general, these studies refer to the overextension of *estar* in contexts were *ser* is the canonical copula. In a diachronic study of the grammaticalization of *ser* and *estar* in Romance languages, Batllori and Roca (forthcoming) argue that the progressive reduction of *ser* and the extension of *estar* is a strategy used by the child since it leads to structural simplification. Therefore it is important to observe whether children of different dialects of Spanish, that is, different inputs, differ in their representations of both copulas.

3.1.3. Experimental methodology

The experimental design that we used was a PST. Each different trial consisted of two pictures, a description of the pictures (to draw equal attention to both pictures) and a experimental question. The question was always asked with copula *estar* but the pictures were different in that one of them depicted a specific unmarked characteristic or property consistent with a *ser* interpretation while the other depicted the same characteristic enhanced by a particular situation, consistent with an *estar* interpretation as illustrated in (2).

(2)   Example of Pictures used in the PST.
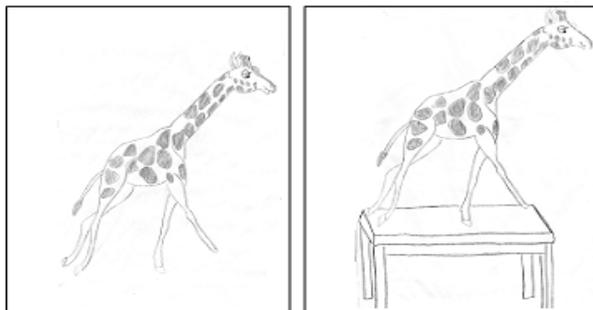
A finding that children behave as adults would suggest that children know that *estar* adds information to a previously known state. In contrast to Schmitt et al.'s (2004) study in which one picture depicted typical properties while the other depicted atypical properties, the current experiment used pictures in which the character retained her normal properties (e.g. a normally tall giraffe on top of a table). That is, we created visual contexts that were consistent with Luján's (1981) account that an *estar* representation entails a *ser* representation. On the contrary, the pictures used by Schmitt et al. (2004) used contexts in which the appropriate pragmatic inferences cancelled the entailment relationship (e.g. Picture A depicted a tall giraffe on the floor consistent with both a *ser* and *estar* tall interpretation while picture B depicted a dwarf giraffe on top of a table consistent with an *estar* tall interpretation). That is, the contexts they used gave rise to the implicature that *estar* + property implies negation of the *ser* + property. The reader is referred to the extensive literature on scalar implicatures for details (Grice, 1989; Chierchia, 2000; Noveck, 2001; Papafragou & Musolino, 2003; Guasti, Chierchia, Crain, Foppolo, Gualmini & Meroni, 2005).

Another important departure from the study of Schmitt et al. (2004) was that we only tested *estar*. This was done because we wanted to directly test whether children overuse *estar* and we also wanted to lower the task demands exerted over

the participants. Since in the current task both pictures are compatible with a *ser* interpretation we simplified the participants' task by using only *estar*. This way we could be more certain that if differences between adults and children are found they are likely to reflect grammatical knowledge.

3.1.3.1. Procedure and materials

Both pictures were presented to the child at the same time and the experimenter made sure the child looked at both pictures before presenting them with the experimental sentence. The child's task was to look at both pictures and select (by pointing) the one that best answered the experimental question. As the objective of the task was to test the comprehension of *estar*, all the experimental sentences asked for a specific property with *estar*. The experimenter took notes of the picture the child pointed to immediately after the child had made the choice. The adults were tested following exactly the same procedure as the one used with children. The experiment was presented to the children individually by one experimenter. The Chilean children were tested in a quiet area of the school that they were attending to in the Punta Arenas region while the Cuban children were tested in a quiet room at their kindergarten in Havana City. The materials were presented in a session of 5 to 10 minutes.

A total of 26 pictures including fillers for Chileans and 24 pictures for Cubans were used in this task. The pictures were presented to children and adults in the same order. Both pictures depicted the same subject but in different situations. Subjects were presented with a brief description of the two pictures. After the story, the child was asked *¿Cuál jirafa está alta?* 'Which giraffe is (*estar*) tall'? Given the context, the expected adult response would be to point to the picture in which the giraffe is on the table. The difference between the pictures is that while the giraffe on the left is intrinsically tall, the giraffe on the right is both intrinsically and extrinsically tall. Thus, although it is true that both giraffes *ser* tall, it is not completely appropriate to say that both giraffes *estar* tall in both pictures. Hence if children know that *estar* is marked, they should select the picture in which the giraffe is on the table because this is the picture that adds more information to the already known state of giraffes being tall animals.

The Cuban PST contained three experimental adjectives, *linda* [pretty], *alta* [tall], *bajita* [short], and a control item with adjective *rubia* [blonde]. The Chilean PST contained four experimental adjectives: *rayada* [listed], *alta* [tall], *bajita* [short], *rojo* [red], and a control item with adjective *rubia* [blonde]. In addition, fillers that asked about the location of an object were included (see Appendix B for all items).

3.1.4. Results

Given that this was a dual choice task, that is, the subject had to select one from two possible pictures, we tested whether the probability of answering correctly (selecting the picture that better matches the experimental sentence) was higher than .5 or at random. The results of the binomial tests show the probabilities of answering correctly significantly above .5 for all the items except for the control item *rubia* [blonde]. Table 1 summarizes the mean

Table 1. Mean percentages correct per adjective in the Cuban and Chilean groups.

Chilean	Adults (n=25)	Children (n=30)
Linda [pretty]	72% (.00)	83.3% (.00)
Alta [tall]	88% (.00)	93.3% (.00)
Bajita [short]	100% (.00)	93.3% (.00)
Rubia [blonde] (control)	52% (.5)	36.6% (.9)
Cuban	Adults (n=9)	Children (n=18)
Rayada [listed]	88.8% (.02)	83.3% (.00)
Alta [tall]	100% (.00)	88.8% (.00)
Bajita [short]	100% (.02)	83.3% (.00)
Rojo [red]	88.8% (.00)	88.8% (.00)
Rubia [blonde] (control)	88.8% (.00)	11.11% (.99)

percentage of correct responses per adjective in both adult and children in the Chilean and Cuban group respectively. The p-values that are given in parentheses are calculated in order to show if the percentage per adjective is significantly above chance.

An independent samples t-test that compared the means of Chilean adults and Chilean children revealed that there are no significant differences between children and adults in the experimental items (Mean = 97, SD = 8.3) and children (Mean = 86.1, SD = 23.04); t (23.6) = -1.82, p = 0.081. Likewise, an independent t-test that compared the performance of Cuban adults (Mean = 90, SD = 15.53) and Cuban children (Mean = 86.7, SD = 19.24) revealed that there are no significant differences with t (53)= 0.711, p= .48.

The data from the Cuban subjects were entered into a one-way independent Anova with mean percentages correct as the dependent variable and age as the independent variable with three levels (young children aged 3;3 to 4;7, old children aged 5;2 to 7;3, and adults) in order to check whether there were any differences related to age. Table 2 shows the descriptive statistics for the Cuban group. The assumption of homogeneity of variances was violated (Levene's Test, p= 0.002). Therefore a Welsch F Test (Anova) was performed indicating that the means of the three age groups were statistically non-significant: F (2, 25.9) = 2.87, p = 0.074.

Likewise, the data form the Chilean subjects were entered into a one-way independent Anova in order to compare the means of the three age groups. Table 3 shows the descriptive statistics.

Again, the assumption of homogeneity of variances was violated (Levene's test, p = 0.00). However, the Welsch Anova could not be performed because the variance of the older group of children was equal to zero. Therefore we used the harmonic mean size that reduces the bias that arises when sample sizes are unequal as in this case (Field, 2009). A Tukey HSD test indicated that the means of the three age groups were not statistically different (p= .07). Nevertheless, even though we do not observe significant differences on the means by age group the results indicate that older children perform better than younger children in both dialectal populations suggesting that the younger children are still progressing in their acquisition of es*tar*.

3.2. The Acceptability Task

3.2.1. Goal of the experiment

Like in Schmitt et al. (2004), the main goal of the AT was to test whether and how Spanish-speaking comprehension of *ser* and *estar* differs in three different conditions: a lexical-semantic condition named Polysemic Free, a syntactic condition named *De* Complement and a discursive-pragmatic condition named Context Only. This study goes one step forward by adjusting the results for a number of factors such as attention and yes-bias, in a unified regression framework. Hence we were able to obtain estimates of the difficulty associated to processing each type of information: lexical-semantic, syntactic, and discursive-pragmatic.

Table 2. Means by age groups in the PST (Cuban sample).

	N	Mean	SD
Young children	11	81.8	17.4
Old children	19	94.7	12.5
Adults	25	86.7	19.24

Table 3. Means by age groups in the PST (Chilean sample).

	N	Mean	SD
Young children	12	79.2	25.7
Old children	6	100	.00
Adults	9	97.2	19.9

3.2.2. Participants

Twenty-seven monolingual Cuban Spanish-speaking children (mean age 62.2 months, range 39- 87 months, SD = 11.8, equivalent to 5;2, range 3;3-7;3) participated in this study. There were four 3-year-olds, five 4-year-olds, thirteen 5-year-olds, four 6-year-olds and one 7-year-old in the sample (13 boys and 14 girls). They had a mixed socioeconomic background with all children attending public kindergartens located in Havana City. In addition, 23 adults were included as controls.

3.2.3. Experimental methodology

The experimental design that we used was an Acceptability Task, a version of the Grammaticality Judgment Task (McDaniel & Cairns, 1996). Normally, children need to judge whether a sentence uttered by an experimenter is well-formed or has a correct interpretation. Every trial consisted of a short story for context-setting purposes and one experimental sentence that matched (or not) the context.

3.2.3.1. Procedure and materials

The task was presented to the children by one experimenter as a game in which the child had to teach Spanish to a puppet called Pepe. The child's task was to listen to a particular short story read by the experimenter and then judge whether the sentence uttered by Pepe was acceptable by saying *sí* (yes) or *no* (no). The child also had to feed the puppet a candy if he gave a right answer or a vitamin if he gave a wrong answer. The adults were given a Paper and Pencil test. The experiment was individually presented to the children at their kindergarten in Havana City and took approximately 15 minutes. Children's acceptability judgments were elicited in three conditions as illustrated in Table 4.

The main goal of Condition 1 was to test whether children know that the use of different copulas correlate with the different meanings of a specific adjective. Therefore the Polysemic Free Condition involved adjectives that have different lexical meanings depending on which copula they are used. In turn, the main goal of Condition 2 was to test

Table 4. Experimental conditions and examples of the Acceptability Task.

Conditions	Story	Exp. Sentence
Cond. 1 Polysemic Free	A. Al perro de Juan hay que llevarlo al veterinario porque se rompió la pata. Juanito bañó y peinó al perro antes de salir para que el veterinario lo encontrara limpio y bonito. B. El perro de Juan entiende todo lo que le dicen. Si le dicen 'siéntate', se sienta. Si le dicen 'dame la pata', la da. Parece que lo tienen muy Bien amaestrado porque siempre obedece.	El perro de Juan es/ está listo. [John's dog is (ser) ready] [John's dog is (estar) clever]
Cond. 2 'De' Comp.	Luis comió mucho en el cumpleaños de su hermana y la comida le cayó mal. Ahora no se siente muy bien y la barriga le pesa como si tuviera piedras.	Luis *es/está muerto de hambre. [Luis is dying of hunger]
Cond. 3 Context Only	A. A María la conocen como la conversadora del Barrio porque generalmente habla mucho pero hoy no ha dicho nada. Todos se extrañan porque nadie le ha podido sacar una palabra. B. Todos conocemos a María desde chiquita y nunca habla mucho. En realidad, casi nunca abre la boca, pero hoy algo extraño le pasa porque no para de hablar. Que tendrá María hoy!	María es/está callada. [M. is a quiet person] [M. is being quiet right now]

whether children know that *estar* but not *ser* takes adjectives with *de* [of] complements. This condition explores a semantic/syntactic restriction in the use of the copula. When an adjective takes a *de* complement, the only grammatical option is to use copula *estar*. Thus, children could rely on syntactic information in order to verify whether the choice of the copula is or not acceptable independent of the story and/or the exact meaning of the adjective. Finally, the main goal of Condition 3 was to test whether children know that *estar* is used with contingent properties ('now') while *ser* is used with more permanent or inherent properties. This condition is very similar to Condition 1 (Polysemic Free) except that here the adjectives do not have different lexical meanings. In Condition 3 copula *ser* is used to express that the property is not bound to the discourse and therefore it qualifies as a permanent property. The copula *estar*, in contrast, is used to describe a temporary property. In order to show adult-like knowledge in this condition, children would have to rely entirely on contextual information.

The predictions for this task are: a) children will perform better in the experimental sentences with *estar*, b) children will make more errors in Condition 3 since copula choice is determined by context alone, and c) children will perform better in Condition 2 since copula choice is determined by syntactic factors.

3.2.4. Results

In this task, the subjects were expected to answer either 'yes' or 'no' after each trial according to the context set by the short story. It is a well known fact in the psycholinguistics literature that respondents tend to use 'yes' as a default answer. That is, when subjects are not sure about what the correct answer is, they tend to respond 'yes' to any question regardless of whether the expected correct response is 'yes' or 'no'. This bias is common in experiments with children (Chien & Wexler,

1990; Schmitt et al., 2004; Waterman, Blades & Spencer, 2004). Therefore, first we provide the means per condition and then we use regression analyses to obtain estimates that are free from the 'yes'-bias phenomenon. Researchers in linguistics normally deal with this problem by eliminating the information for items that have an affirmative answer as the correct response (Schmitt et al., 2004). Although this procedure effectively leaves only observations that are bias-free, this is done at the cost of eliminating substantial information, with the consequent loss in estimating precision. Hence, it is important to note that the results vary depending on the statistics used. Regression analyses are extremely useful for controlling extra variables such as attention, experimental noise and response biases that can affect subjects' performance. Since a major objective of this experiment is to explore whether the experimental variables differentially predict outcomes on *ser/estar*, the use of regression is ideal since it allows us to rank these factors in order of processing difficulty. The general means of the subjects in the three conditions are presented in Table 5.

A Chi-square test shows that the difference between the means of adults and children is only significant in Condition 3, with χ^2 (1, N= 38) = 6,01, p= .014. In the other conditions the difference between children and adults was not significant (p>.05). A Chi-square test also shows that the means of condition 1, 2 and 3 for children are not significantly different (p> .05).

Now, when we refer to the regression analysis the results change. Table 6 shows the variables recorded in the database, and Table 7 shows the descriptive statistics.

As said previously respondents tend to use 'yes' as a default answer even when the appropriate answer is 'no'. Therefore, we begin the analysis by testing for a possible 'yes'-bias. Because it is observed in children more than in adults, it leads to higher differences in observed performance between adults and children than those that effectively exist.

Table 5. Means per condition in adults and children.

Condition	Adults	n	Children	n
Cond 1	0.8867	22	0.6245	16
Cond 2	0.9848	22	0.7751	16
Cond 3	0.8986	23	0.5796	20

Table 6. Variable description for the regression analysis.

Variable Name	Description
Mainvar	=1 if answer is correct, =0 otherwise
Child	=1 if respondent is a child, =0 otherwise
Age	Age, in months. Recorded only for children
Dyes	=1 if correct answer is 'yes', =0 if correct answer is 'no'
Filler	Percentage of correct answers in filler questions

Table 7. Descriptive statistics.

Variable	Mean	SD	Min.	Max.	No obs.
Mainvar	74	.44	0	1	1135
Child	.46	.5	0	1	49
Age	62	11.7	39	87	27
Dyes	.5	.5	0	1	1135
Filler	.84	.14	.43	1	50

A probabilistic analysis indicates that the probability of adults for answering yes/no is exactly .5 so no bias is observed in the control group. However, this probability is .23 for children, with the hypothesis of no bias rejected at the 1% confidence level.

Once the bias is established, the question remains of how to obtain estimates that are free from it. The first regression estimates the probability of giving a correct answer controlling for the 'yes' bias. The results indicate that the probability of answering correctly for adults (child=0) when there is no bias (dyes = 0, dyes*child = 0) is .84. For children (child=1), this probability is .84 - .41 x 1= .43. The difference in probabilities of answering correctly between adults and children is statistically significant (p =.00). It is important to notice that the probability that adults answer correctly is less than 100%. As we are comparing child knowledge of *ser* and *estar* with respect to the adult grammar, we focus on the difference in performance between children and adults.

We will argue that the relevant statistic to focus on is the difference in probabilities between adults and children. We assume that the failure of adults answering correctly is due to the fact that the answers reflect not only understanding of the subject on *ser* and *estar* but also the quality of experimental design (See the empirical model in Appendix D).

We proceed to report the results regarding the following questions: i) Do children show different performance between the copulas that would allow us to suggest that *estar* is acquired before *ser*?, ii) What kind of information (syntactic, semantic or pragmatic) do children use when distinguishing *ser* and *estar*, and iii) Do we observe age effects?

In order to answer the first question, we examined whether children show better performance with *estar* than with *ser*. As *estar* can be a subset of *ser*, it is reasonable to expect that children acquire *estar* before *ser*. To test this hypothesis a variable that captures the difference between children and adults in *ser/estar* comprehension was created (*ser*child). The result indicated that children have less trouble with *estar*, but the variable's coefficient was not statistically significant.

In order to answer our second question that asked in which order relative to difficulty our three experimental conditions (syntactic, semantic or pragmatic information) are used by children when employing *ser* and *estar*, we used the difference in probability between children and adults (or b1), as the object of interest (Table 8).

A *t*-test for equality of means shows that coefficient (in condition) 3 is statistically higher than 1 (p=.00), and that coefficient 1 is statistically higher than

Table 8. Coefficient b1 (p-values in parentheses).

Condition	Coefficient	SD	P-value	Statistical ranking
1	-.42	.0778	(.00)	2
2	-.33	.1023	(.00)	1
3	-.48	.0648	(.00)	3

2 (p=.00). These results allow us to say that for children it is easier to use syntactic information in order to choose between *ser* and *estar*. In other words, children performed better in Condition 2 (*de* Complement) perhaps because the presence of a *de* complement helps them to select *estar* over *ser*.

The results also indicate that children seem to perform better when they have lexical/semantic cues than when they have only contextual cues. Notice that these results are different from the results we obtained by comparing means. When considering the means children and adults performance was statistically different only in the 'Context Only' condition. Here, we found that children and adults perform differently in all conditions and that the difficulty across conditions is clearly ranked.

In order to answer our third question, we tested whether the comprehension of *ser* versus *estar* depends on the age of children, but again the coefficient associated with the interaction *ser/ estar* x age was found to be small and statistically insignificant. We also found a significant positive correlation between the age of children and their percentage of correct responses in the filler questions (p=.00), showing that children do pay more attention as they get older. This should of course be of some concern: it could be that all the differences in correct responses in the main questions are due simply to the fact that adults pay more attention to the questions themselves. To address this issue, we did the previous analysis again controlling for the attention factor summarized in the filler variable. We found that the differences between adults and children become smaller, so the attention effect exists, but is quantitatively very small. With the inclusion of the filler variable, the coefficient of interest (on the variable *child*) is -.38 instead of -.41, a difference of 3 percentage points. In general, the results strongly confirm that the differences in performance are not driven by differences in attention. Finally, it is worth pointing out that we found significant age effects in the fillers

but *not* in answers to the experimental sentences, again casting doubt on a simple explanation of performance driven by the capacity of concentrating of older versus younger children.

3.3. Discussion of the PST and the AT

Our main finding is that children behave adult-like in the PST but differed from adults in the AT. This is not in itself a contradiction since the methodologies used are different. In the PST the child makes a choice between two alternative pictures, while in the AT the child is asked to evaluate the 'Acceptance' of either a correct/felicitous or an incorrect/unfelicitous linguistic string according to a specific context set by the story. The AT is much more demanding than the PST because while in the PST the child selects a preferred option by directly accessing the visual context in the AT the she/he needs to remember the story and also compare the two competing representations (e.g. the correct and the incorrect) in order to evaluate whether the sentence provided matches or not the interpretation of the story. Thus it is expected that children will have more problems when copula knowledge is tested by methods that rely on processing capacities (Grodzinsky & Reinhart, 1993). Our line of reasoning is supported by previous studies (Avrutin & Wexler, 1992) that show that children perform poorly when pragmatics is involved. Similarly in the AT children performed least adult-like when the choice of copula was determined solely by context. Therefore, it cannot be argued that Spanish-speaking children do not distinguish between *ser* and *estar*.

Having this in mind we turn to discuss the results of each task separately. The main objective of the PST was to explore whether Cuban and Chilean Spanish speaking children know that *estar* is the discursively linked copula. It was predicted that if children know the discursive-pragmatic properties of *ser* and *estar* they should consistently select the picture that is more appropriate to match an *estar* interpretation.

The first important finding of the PST is that children show an adult-like performance in the understanding of *estar*. This result suggests that they know that *estar* relates to a specific situation. Crucially, the different performance between the control (the blonde trial) and the experimental items indicates that subjects experience more difficulty when the context requires to draw the pragmatic inference that *estar* is used in order to negate *ser*. That is, even though *ser* and *estar* are partially synonyms (Luján, 1981) they can be used to make pragmatically different inferences. For example, by saying *Marta está linda hoy* we may implicate that Marta is not a pretty person. In this task both Cuban adults and children selected the typically blonde girl consistent with a *ser* blonde interpretation 48% and 63% of the time respectively (chance behavior) while Chilean children selected the *ser* picture 89% of the time in clear contrast to Chilean adults who only selected the *ser* picture 11% of the time. Therefore, processing the inferences associated with *estar* appears to be more difficult in contexts like the one used in the control item (blonde) than in the contexts used for the experimental items. A possible explanation is that the context in the control item but not in the experimental items gives rise to the pragmatic inference that *estar* negates *ser*. This result falls in line with most studies on the acquisition of pragmatic inferences (Noveck, 2001; Papafragou & Musolino, 2003; Chierchia et al., 2004) that processing the appropriate computation of referential elements (i.e. inferences) in specific contexts is costly (Grodzinsky & Reinhart, 1993; Reinhart, 2006). That children could draw the *estar* inferences in the experimental items but not in the control item (as in Schmitt et al., 2004) is also consistent with Reinhart's (2006) explanation about children's failure to compute pragmatic inferences. Children may be aware of the required computation, but unlike adults, they are unable to execute it because they have limited processing resources. Since the context of the control item was more difficult to process than the context in the experimental items (for both children and adults), children were able to compute the inferences associated with *estar* in the less demanding contexts (experimental) but failed to do so in the more complex one (control).

Regarding the control item, we also need to explain why Cubans and Chileans vary in their performance. When asked *¿Cuál niña está rubia?*, Cuban adults and children selected the picture consistent with a *ser* interpretation 48% and 63% of the time respectively while Chilean adults and children did so 11% and 89% of the time respectively. That is, Chilean adults are more restrictive than their Cuban counterparts for allowing typically *ser* properties to be described with *estar*. This result may be related to the process of *estar* overextension that has been described in the sociolinguistic literature and could indicate that the encroaching of *estar* into *ser* territory is more common in Cuban than in Chilean Spanish.

The second finding of the PST is that children do not seem to use typicality cues of adjectives since they assigned *estar* interpretations to adjectives that are typically used with *ser*. Hence, we do not have evidence for Sera's (1992) claim that children acquire adjectives with either a *ser* or and *estar* status.

Regarding our second task, the AT, our first finding was that children did not perform adult-like in any condition. However, the results obtained allowed us to rank children's performance in the following manner in the order of most to least adult-like: (1) *De*-Complement, (2) Polysemic Free, and (3) Context Only. Since the difference in performance between the conditions was found to be statistically significant and because every condition controlled for different properties of *ser* and *estar* we can interpret this ranking as showing the kinds of knowledge children exploit in order to use copulas *ser* and *estar*. The result that children performed better in condition 2, where they had to rely on the syntactic requirement that *estar* but not *ser* can take *de* complements, confirms Sera's (1992) intuition that Spanish-speaking children may use syntactic cues in order to discriminate between the copulas. The fact that children's performance was better in this condition than on the condition that controlled for lexical meaning of the frame Copula+Adjective, suggests that children rely on syntactic information more than in semantic information. Also, the poor performance of children in the condition that controlled for contextual factors alone, supports our prediction that children have more problems with the pragmatic dimension of copular use.

It was also found that children performed better with *estar* than with *ser*. However, the difference between children's performances with *ser* and *estar* were not statistically significant and therefore, we could not confirm the hypothesis that *estar* is acquired earlier than *ser*.

At last but not least, we found that age did not affect children's performance in the experimental sentences. It is important to note, however, that older subjects performed better than younger subjects in the fillers. We interpret this as showing that the experimental results that we obtained were not highly dependent on attention or other extra-linguistic factors that could interfere with subject's performance. Therefore, we consider the differences in performance to reflect grammatical competence. We also found that attention has a statistically significant role in explaining differences between adults and children. That is, as children pay less attention than adults, part of the difference between these two age groups corresponds purely to differences in attention and not necessarily to different representations of *ser* and *estar*. Therefore, an advantage of the current study is that we could isolate the effect of *ser* and *estar* comprehension.

CONCLUSIONS

We conducted two comprehension experiments of *ser* and *estar* in Spanish child language. The first experiment was a PST that tested comprehension of the discourse properties of *estar* and the second experiment was an AT that tested syntactic, semantic and pragmatic properties of *ser* and *estar*. The results expanded the findings of Schmitt and Miller (2007) by showing that Spanish-speaking children know that *ser* and *estar* have different discourse properties. Unlike Schmitt et al. (2004) but similarly to Schmitt and Miller (2007), it was found that Spanish speaking children interpret the pragmatics associated to *estar* just as adults do. Another important finding was related to the question of whether children divide adjectives as having a *ser* or *estar* status as motivated by Sera's (1992) claim that Spanish-speaking children draw a distinction between *ser* adjectives and *estar* adjectives. Given that in the PST children assigned an *estar* interpretation to adjectives that typically take *ser*, it is unlikely that children acquire adjectives by distinguishing those that typically appear with *ser* from those that typically appear with *estar*. The results of the AT, in turn, are consistent with Schmitt et al.'s (2004) finding that the pragmatic dimension of copula use is still complex for 3- to 7- year-old children. The fact that children performed much

better in the condition in which *estar* (but not *ser*) takes *de* complements than in the conditions that controlled for lexical/semantic factors, suggests that children are more sensitive to distributional cues than to lexical/semantic cues. Moreover, as children performance was least adult-like in the condition in which the only clues for copula usage were given by the context, we argued that children have more problems with discourse factors. Therefore, in the order of more used information to the least used information, we can arrange them as follows: first, syntactic, second lexical/semantic and third contextual information alone.

An important question dealt with the issue of which copula is acquired earlier in Spanish language. Although we saw a tendency in the right direction, namely that children performed better with *estar* sentences, the results lacked statistical significance. However, it is still possible that children use the Semantic Subset Principle in early stages of copula acquisition. Further research will need to address if Spanish-speaking children use the Semantic Subset Principle for distinguishing between *ser* and *estar*.

Finally, contrary to Schmitt et al. (2004) and Sera (1992) the results do not lend support to the claim that *estar* is overused. In the PST the children performed adult-like when assigning the pragmatically appropriate interpretation to the experimental questions with *estar* but were completely different (at chance) in the control item that was similar to the experimental items used in Schmitt et al. (2004). These results suggest that the pictures used in the PMT of Schmitt et al. (2004) (and our control item) were contextually more complex and harder to process. Therefore the *estar* overuse observed in Schmitt et al. (2004) is more likely to reflect a problem of limited processing resources by children rather than a problem for drawing the right inferences for *estar*. Similarly, the fact that in the AT children performed better with *estar* than with *ser* is clearly incompatible with an overuse of *estar* by children. The current results highlight the need for further investigating which copula (if any) is overused in the acquisition of Spanish by monolingual children.

REFERENCES

Arche, M. (2006). *Individuals in time; tense, aspect and the individual/stage distinction.* Amsterdam & Philadelphia: John Benjamins.

Avrutin, S. & Wexler, K. (1992). Development of Principle B in Russian: Coindexation at LF and Coreference. *Language Acquisition: A Journal of Developmental Linguistics*, *2*(4), 259-306.

Batllori, M. & Roca, F. (forthcoming). Grammaticalization of SER and ESTAR in Romance. In D. Jonas, J. Whitman & A. Garrett (Eds.), *Grammatical Change: Origins, nature, outcomes.* Oxford: Oxford University Press.

Bello, A. (1951). *Gramática de la lengua Castellana.* Buenos Aires: Anaconda.

Clements, J. C. (1988). The semantics and pragmatics of the Spanish <Copula+Adjective> construction. *Linguistics, 26*(5), 779-822.

Clements, J. C. (2005). 'Ser' and 'estar' in the predicate adjective construction. In J. C. Clements & J. Yoon (Eds.), *Functional approaches to Spanish syntax: Lexical semantics, discourse, and transitivity* (pp. 161–202). London: Palgrave-Macmillan.

Crain, S. & Thornton, R. (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics.* Cambridge, MA: MIT Press.

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.

Chierchia, G. (2000). *Scalar implicatures and polarity items.* Paper presented at NELS 31, Georgetown University, Washington, DC.

De Jonge, B. (1993). (Dis)continuity in language change: Ser and estar + age in Latin American Spanish. In F. Drijkoningen & K. Hengeveld (Eds.), *Linguistics in the Netherlands* (pp. 69-80). Amsterdam: Benjamins.

Diesing, M. (1990). Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory, 8*(1), 41-79.

Diesing, M. (1992). *Indefinites.* Cambridge, MA: MIT Press.

Fernández, M. (1999). La predicación: Las oraciones copulativas. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 2354-2460). Madrid: Espasa.

Field, A. (2009). *Discovering statistics using SPSS.* London: Sage publications.

Gili Gaya, S. (1955). *Curso superior de sintaxis española.* Barcelona: Spes.

Grice, H. (1989). *Studies in the ways of words.* Harvard: Harvard University Press.

Grodzinsky, Y. & Reinhart, T. (1993). The innateness of binding and of coreference. *Linguistic Inquiry, 24*(1), 69-101.

Guasti, M. T., Chierchia, G., Crain, S., Foppolo, F., Gualmini, A. & Meroni. L. (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*(5), 667-696.

Gutiérrez, M. (1992). The extension of estar: A linguistic change in progress in the Spanish of Morelia, Mexico. *Hispanic Linguistics, 5*(1-2), 109-141.

Holtheuer, C. (2009). *Learning Ser and Estar: A Study of production and comprehension in Chilean Spanish.* PhD thesis, Australian National University, Australia.

Luján, M. (1981). The Spanish copulas as aspectual indicators. *Lingua, 54*(2-3), 165-210.

Maienborn, C. (2003). Against a Davidsonian analysis of copula sentences. In M. Kadowaki & S. Kawahara (Eds.), *NELS 33 Proceedings of the Thirty-Third Annual Meeting of the North East Lingusitic Society* (pp. 167-186). Amherst, MA: Create Space.

Maienborn, C. (2005). A discourse-based account of Spanish ser/estar. *Linguistics, 43*(1), 155-180.

McDaniel, D. & Cairns, H. (1996). Eliciting judgments of grammaticality and reference. In D. McDaniel, C. McKee & H. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 233-254). Cambridge, MA: MIT Press.

Noveck, I. A. (2001). When children are more logical than adults: Investigations of scalar implicature. *Cognition, 78*(2), 165-188.

Papafragou, A. & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics pragmatics interface. *Cognition, 86*(3), 253-282.

Reinhart, T. (2006). *Interface strategies: Optimal and costly computations.* Cambridge, MA: MIT Press.

Schmitt, C. (1992). Ser and estar: A matter of aspect. In K. Broderick (Ed.), *NELS 22 Proceedings of the Northeastern Linguistic Society* (pp. 411-426). Amherst, MA: GLSA.

Schmitt, C. (1996). *Aspect and the syntax of noun phrases.* PhD Dissertation. University of Maryland, USA.

Schmitt, C. (2005). Semi-copulas: Event and aspectual composition. In P. Kempchinsky & R. Slabakova (Eds.), *Syntax, semantics and the acquisition of aspect* (pp.121-145). Springer: Kluwer.

Schmitt, C., Holtheuer, C. & Miller, K. (2004). Acquisition of copulas ser and estar in Spanish: Learning lexico-semantics, syntax and discourse [online]. Retrieved from: http://www.bu.edu/linguistics/APPLIED?BUCLD/supp.html

Schmitt, C. & Miller, K. (2007). Making discourse-dependent decisions: The case of the copulas ser and estar in Spanish. *Lingua, 117*(11), 1907-1929.

Sera, M. (1992). To be or to be. *Journal of Memory and Language, 31*(3), 408-427.

Sera, M., Bales, D. & del Castillo, J. (1997). Ser helps speakers identify "real" properties. *Child Development, 68*(5), 80-831.

Silva-Corvalán, C. (1991). Spanish attrition in a contact situation with English. In H. Seliger & R. Vago (Eds.), *First language attrition* (pp.151-171). Cambridge: Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change.* Oxford: Oxford University Press.

Silva-Corvalán, C. & Montanari, S. (2008). The acquisition of ser, estar (and be) by a Spanish-English bilingual child: The early stages. *Bilingualism: Language and Cognition, 11*(3), 341-360.

Vañó-Cerdá, A. (1982). *Ser y estar + adjetivo. Un estudio sincrónico y diacrónico.* Tübingen: Verlag.

Waterman, A., Blades, M. & Spencer, C. (2004). Indicating when you do not know the answer: The effect of question format and interviewer knowledge on children's "don't know" responses. *British Journal of Developmental Psychology, 22*(3), 335–348.

APPENDIX

Appendix A. Norming Study for Picture Selection Task

Fill in the blanks with either *ser* or *estar*.
1)   *La sirenita ___ linda*
2)   *Juanito ___ listo*
3)   *Lola, la jirafa ___ alta*
4)   *La comida ___ caliente*
5)   *La pulgarcita ___ bajita*
6)   *La niña ___ rubia*
7)   *El plátano ___ Amarillo*
8)   *El plato ___ vacío*
9)   *La cebra ___ rayada*
10)   *El tomate ___ rojo*
11)   *La película ___ buena*
12)   *Pedro ___ cansado*

Appendix B. Items for the PST

	Cuban group	**Chilean group**
Experimental items	*¿Cuál sirenita está linda?*	*¿Cuál cebra está rayada?*
	¿Cuál jirafa está alta?	*¿Cuál jirafa está alta?*
	¿Cuál pulgarcita está bajita?	*¿Cuál pulgarcita está bajita?*
	¿Cuál niña está rubia? (Control)	*¿Cuál tomate está rojo?*
		¿Cuál niña está rubia? (Control)
Fillers	*¿Cuál televisor tiene una silla?*	*¿Cuál silla está sobre un televisor?*
	¿Cuál maleta tiene un barco?	*¿Cuál barco está sobre una maleta?*
	¿Cuál botella tiene una revista?	*¿Cuál revista está sobre una botella?*
	¿Cuál sombrero tiene un zapato?	*¿Cuál sombrero tiene un zapato?*
	¿En cuál taza hay un plato?	*¿En cuál taza hay un plato?*
	¿En cuál carro hay una bicicleta?	*¿En cuál auto hay una bicicleta?*
	¿Cuál jirafa patea a la mujer?	*¿Cuál jirafa patea a la mujer?*
	¿En cuál mesa hay un camión?	*¿En cuál mesa hay un camión?*

Appendix C. Items for the Acceptability Task

Condition	**Experimental Items**
Condition 1 Polysemic Free	*La radio es/está buena.*
	El perro de Juan es/está listo.
	El gato es/está vivo.
	La lavadora es/está mala.
Condition 2 'De' Complement	*Pedro *es/está mal de la barriga.*
	*Juan *es/está aburrido de ver televisión.*
	*Pepa *es/está orgullosa de sus hijas.*
	*Luis *es/está muerto de hambre.*
Condition 3 Context Only	*María es/está muy callada.*
	Pepa es/está blanca.
	Claudia es/está colorada.
	Los ojos de Juan son/están rojos.

Appendix D. General empirical model for the probability that adults and children answer correctly in the AT

(1) Prob (correct | adult) = 1+ b1 Ex design + b2 (other factors)+e1 (adults)
(2) Prob (correct | child)= a + b1 Ex design + b2 (other factors) + e2 (children)
(3) Prob (correct | Adult) – Prob (correct| child) = 1-a + (b1 Ex design -b1 Ex design) + (b2 (other factors) - b2 (other factors)) + e1-e2 = 1 - a + e3 (adults- children)