UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRIA Y DOCTORADO EN PSICOLOGIA
ANÁLISIS EXPERIMENTAL DEL COMPORTAMIENTO (AEC)

LA ADQUISICIÓN DE LA COMPRENSIÓN DEL FOCO PROSÓDICO EN HABLANTES DEL ESPAÑOL: DE LA PERCEPCIÓN AL ANÁLISIS

TESIS
PARA OPTAR POR EL GRADO DE:
DOCTOR EN PSICOLOGIA

PRESENTA:
BRENDA MUÑOZ MARTÍNEZ

TUTOR PRINCIPAL:
DRA. NATALIA ARIAS TREJO
FACULTAD DE PSICOLOGÍA, UNAM

MIEMBROS DEL COMITÉ TUTOR:

DRA. ALEJANDRA AUZA BENAVIDES
HOSPITAL GENERAL DR. MANUEL GEA GONZÁLEZ, MÉXICO

DRA. KRISZTINA SZENDRÖI
UNIVERSIDAD DE VIENA, AUSTRIA

DR. JONATHAN AZAEL CABALLERO MENESES
THE CENTER FOR IMPLEMENTATION, CANADÁ

DRA. ERIKA MENDOZA VÁZQUEZ
INSTITUTO DE INVESTIGACIONES FILOLÓGICAS, UNAM, MÉXICO

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO, MARZO 2024.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (Conacyt) por la beca (clave 2019-000037-02NACF) otorgada durante la realización de este proyecto de doctorado (CVU 744095).

A mi comité tutor: a la Dra. Natalia Arias Trejo, por su guía y atenciones durante la realización de este proyecto y durante mi estancia en el laboratorio; a la Dra. Alejandra Auza Benavides, a la Dra. Krisztina Szendröi y al Dr. Jonathan Azael Caballero Meneses, por su retroalimentación y apoyo en cada paso del proceso; a la Dra. Erika Mendoza Vázquez, por sus enseñanzas e interés constante en el proyecto, motivos por los cuales se sortearon dificultades y fue posible esta investigación.

A la Lic. María Fernanda García Silva García y al Mtro. Luis Roberto Bolaños Godoy, por el apoyo en la realización de los estímulos experimentales.

Al Dr. Stanislav Mulík, por la revisión de los análisis de rastreo ocular.

A Luis Muñoz Martínez, por las lecturas de esta tesis que terminaron en su versión definitiva.

A mis compañeros del Laboratorio de Psicolingüística: a sus lecturas y recomendaciones, sin las cuales el proyecto no hubiera adquirido forma y coherencia; además de su apoyo constante durante mi estancia.

Finalmente, a los padres y a los propios participantes de esta investigación. Gracias por su interés, por su paciencia y por su tiempo, sin los cuales no existiría este proyecto.

Dedicatoria

Dedico este trabajo de investigación a mis padres, Lucila y Luis: la razón por la cual me fue posible contar con una educación no sólo académica, sino humana. Gracias por la biblioteca mágica, por las tardes ayudándome con la tarea sin importar el cansancio; por fomentarme el espíritu científico y la necesidad de, no sólo aprender, sino de prepararnos para servirle a los

demás. Sin su ejemplo, amor y apoyo incondicional, reflejo del espíritu de mis abuelos, no hubiera llegado hasta aquí.

A Roberto, mi esposo, por ser el sostén de nuestro hogar durante la realización de mi doctorado. Sin ti, tu comprensión, contención y respaldo constante no hubiera podido seguir avanzando. Fuiste el pilar de estos cuatro años. Esta tesis es el producto también de tu trabajo.

A mi hermano, Luis, por su paciencia; por las lecturas de la tesis, por escuchar mis prácticas de exámenes. Por ser mi confidente y consejero ante las dificultades. Por ser mi ejemplo para salir adelante.

Y, finalmente, a mis compañeros de doctorado; en especial a Ale y Stan, quienes estuvieron muy presentes en esta última parte del proceso. A los tesistas y prestadores de servicio social del laboratorio, con quienes compartí tantas horas de trabajo en ese cubículo del sótano. A todos ustedes, por los cafés, las risas y confidencias; sobre todo, por el apoyo emocional que no nos dejó solos en el camino.

INDEX

Resumen

El foco prosódico es la información relevante de un enunciado, marcada con prominencia acústica. Puede ser colocado en diferentes posiciones del enunciado; en español, usualmente al final. Estudios en inglés, alemán, holandés y francés sugieren que los niños lo comprenden desde los 3 años (Szendröi et al., 2018; Höhle, 2016); aunque cometen errores, lo cual podría deberse a un control inhibitorio poco desarrollado (Höhle et al., 2016). El presente estudio analizará la comprensión del foco prosódico por niños preescolares hablantes del español de 4 y 5 años (edad en la que su control inhibitorio sigue desarrollándose). Se espera que respuestas verbales u oculares incorrectas en tareas de foco prosódico estarán relacionadas con puntuaciones bajas en tarea de control inhibitorio. Dos experimentos analizarán la comprensión del foco prosódico en el sujeto (incio) o atributo locativo (final) de la oración (variables independientes). En el Experimento 1, los niños observan imágenes de seis objetos animados o inanimados; si el experimentador enfatiza el sujeto incorrecto, los participantes deben corregir el sujeto; si se enfatiza el atributo locativo, deben corregir el atributo locativo. La variable dependiente es el número de respuestas correctas. El Experimento 2 es idéntico, pero utiliza un rastreador visual. Las variables dependientes son los patrones y fijación de mirada. Se evalúa el control inhibitorio en una Tarea de Flancos.

Abstract

Prosodic focus is the relevant information of an utterance, marked with acoustic prominence. It can be placed in other positions of an utterance; in Spanish, usually at the end of the phrase. Studies in English, German, Dutch and French suggest that pre-schoolers understand it since the age of 3 (Szendröi et al., 2018; Höhle, 2016); although they make mistakes, which could be due to underdeveloped inhibitory control (Höhle et al., 2016). The present study analyses the comprehension of prosodic focus by 4-year-old Mexican Spanish speakers (age in which their inhibitory control is still under development). It is expected that incorrect verbal or ocular responses in prosodic focus tasks will be related to low scores in the inhibitory control task. For this purpose, two experiments will analyse the comprehension of the prosodic focus on the subject or prepositional predicative of the sentence (independent variables). In Experiment 1, pre-schoolers look at pictures of six animate or inanimate objects; if the experimenter emphasizes an incorrect subject, the participants must correct the subject; if the prepositional predicative is emphasized, they must correct the prepositional predicative. The dependent variable is the number of correct answers. Experiment 2 is identical but uses an eye-tracker. The dependent variables are patterns and gaze fixation. Inhibitory control is evaluated in a Flanker Task.

1. Introduction to prosodic focus comprehension

Several questions remain about the way children develop certain language abilities. In this sense, this dissertation aims to analyse how infants understand prosody –a category that involves aspects such as intonation, rhythm, and the volume of a sentence we orally produce–. Prosody affects every discourse's interpretation and mediates the first stages of language development. One particular aspect of language that involves prosody, which is the main theme of this research, is "prosodic focus", the word or part of an utterance that is prosodically prominent with respect to the other elements of the utterance because of a greater pitch, vocal duration or volume of the voice (Dorta-Luis, 2008). This word or part of the utterance is emphasised with different purposes, depending on what the speaker wants to communicate.

It is still unclear how prosodic focus comprehension is acquired and developed in different languages. So far, research on the age of acquisition has been contradictory, and experiments designed to assess prosodic focus comprehension have not been completely successful. Recent investigations in languages like English, German, French and Danish (Szendröi et al., 2018; Höhle, 2016) have started to shed a light into new methods used to assess prosodic focus. These new investigations point out that children from these languages understand prosodic focus since about 3 years of age, although some still have difficulties, which may be related with other cognitive processes. Because of this, it is still unclear how prosodic focus acquisition develops in relation to such cognitive processes; moreover, how this acquisition works in languages like Spanish has not been thoroughly studied. The present dissertation investigates these matters.

For this purpose, the definitions of "prosody", "prosodic focus", "informative focus", "contrastive focus", "comprehension", "active verbal memory" and "inhibitory control" will be

explained in the following sections of this chapter. Subsequently, a literature review will be presented.

1.1. Prosody as a Key Factor in Language Acquisition

Prosody is a category that refers to elements such as intonation, accent, tone, pauses, rhythm, and the volume of an utterance (which is a spoken sentence or phrase) (Pynte, 1998, p.79). In a communicative context, it is essential to understand any conversation (Baeza-Álvarez & Rodríguez-Maldonado, 2011, p. 6). Prosody not only shows the attitude, emotions or intentions of the speaker, but it is also responsible for speech segmentation of words inside phrases, for determining the type of utterance (question, affirmation or command), and for signalling the relevant part of the discourse (Perrone-Bertolotti, et al., 2013, p. 2575). It is also crucial for language acquisition.

Different studies have shown that infants distinguish prosodic elements even before birth. Foetuses discriminate their mother's voice from other voices (Fifer & Moon, 1994; Kisilevsky et al., 2003; Kisilevsky et al., 2009), different rhythms present in the prosody of different languages (Nazi et al.,1998), and different sound frequencies (which are not exclusively related to prosody but show that children can distinguish these elements of speech and sound in general) (Shabidullah & Hepper, 1994). This prenatal exposition helps babies to distinguish the sounds of their mother tongue shortly after being born (Moon et al., 2012). Furthermore, prosody is also the first element used for communication, through cries and babbling (Kuhl et al., 2008; Tonkova-Yampol'skaya, 1973).

Aside from being the first aspect of language that children distinguish and produce, prosody also seems to be what guides the acquisition of other elements of language. Namely, according to Goswami (2008, p.148), through child-directed speech (the way adults talk to babies

characterised by an exaggerated prosody and slow speech), prosody helps babies identify the boundaries between words in a stream of speech, which ultimately helps them learn words and their meaning. This ability, called "word segmentation", is developed between 6 and 8 months of age, and prosody is one of the most relevant tools for its development, according to Johnson and Jusczyk (2001). In conclusion, to understand how a complex process such as first language acquisition occurs, it is necessary to study prosody, which guides the first steps of this process.

Prosody is a category formed by different elements with distinct communicative functions depending on the context of the conversation. For instance, according to Quilis (1993, pp. 384-490), through the tone and volume of the voice of the speaker, intonation helps distinguishing commands from statements or questions; it may express the emotions or intentions of the speaker; it is used to create emphasis, and it integrates the words into a coherent discourse (without intonation, speaking would be just a list of words without any relation to each other).

In consequence, prosody has several functions, which are dependant of a given context and may also be interpreted differently by the speaker and the listener. The present study analyses an element of prosody whose function is to mark, through intonation, an emphasis on an element of the spoken sentence. This element, known as *prosodic focus*, will be defined below.

1.1.1. Prosodic focus

Focus is usually defined as the new information on an utterance. This definition considers a conversation in which the speaker has an idea of what the hearer knows, and hence they emphasise the information the hearer does not know or what they want the hearer to pay attention to (Dorta-Luis, 2008).

In this sense, utterances are divided in two parts: *topic*, the part of the utterance that contains known information, and *focus*, the new or relevant information in a given context (Vallduví &

Engdahl, 1996, p. 3). These definitions apply even in written language. When spoken, focus is marked with prosodic elements such as stress on a word or group of words (Ladd, 2008, p. 44; Vallduví & Engdahl, 1996, p. 5), and hence is called *prosodic focus*. The following sentences in Spanish from Mora-Bustos (2010), as well as their English translations, serve as an example (imagining they are part of a conversation in which questions were asked before):

(3) a. *Pedro compró el PERIÓDICO.*

(Pedro bought the NEWSPAPER).

b. *El periódico fue comprado por PEDRO.*

(The newspaper was bought by PEDRO).

In this case, (1a) would answer the question: "What did Pedro buy?" The new information, and hence the focus, would be, in consequence, "THE NEWSPAPER."[1] In contrast, (1b) would answer to the question: "Who bought the newspaper?" The answer, and hence the unknown information so far, would be "PEDRO." It is important to point out that prosodic focus is always present in everyday conversation, and it is not always an answer to a real question, but sometimes an implied one in a particular context (Dorta-Luis, 2008).

This definition of focus, however, is limited in the sense that prosodic focus is not always related to new information. Sometimes, the speaker emphasises a word or set of words depending on their intentions in the conversation. Because of this, prosodic focus may be better defined as the word or set of words in an utterance that are emphasised, compared to the other elements of the utterance, with a higher volume or tone, or even a prolonged pronunciation of the

[1] Capital letters will be used as an indication of the focused or stressed word in this example and the followings.

vocals of the focused element, with different communicative intentions. Particularly, the emphasised word or set of words are the part of the conversation that is not previously assumed by the hearer (Dorta-Luis, 2008). In this sense, prosodic focus is the most important information of a linguistic expression in a communicative context that the speaker wishes to integrate into the pragmatic information of the hearer by adding or substituting information (Dik, 1997, p. 326)

When prosodic focus is used specifically to highlight new information that the hearer does not know, it is defined as *informative focus* (Dik, 1997, pp. 332–334). This definition has to do with its pragmatic meaning. Informative focus is relevant in language acquisition, as it helps children to pay attention to this new or important information. In Spanish, as studied in Mexico City by Villalobos-Pedroza (2019), this occurs through an articulatory effort adults give to the word or set of words they want the child to pay attention to. They do this with an early elevation of the intonational peak before the nuclear syllable ends, as well as an extended duration of this syllable. There is a greater articulatory effort when the focused word is not accompanied by deictics or attentional words such as “Look!”, for example.

Furthermore, children as young as 2 have been shown to understand the relation between tonal prominence and novelty (informative focus). They use this information to learn novel words, as informative focus creates joint attention between the child and the speaker to something new referred to in the discourse (Grassmann & Tomasello 2007; 2010). This means that, from early on in development, informative focus is the most frequently understood type of focus.

To understand how focus is marked prosodically, as well as the different ways Spanish speakers mark focus, the following sections will first explain the importance of word order in this language, and then the prosodic elements that create stress or emphasis on the focused word.

1.1.2. Word order and its relevance for prosodic focus in Spanish

It is relevant to mention that several studies about focus acquisition have been done in English and other languages, like Dutch and German (Szendröi et al., 2018; Höhle, 2016), that mark focus in an equivalent manner to English. For this reason, in the following sections, English will be used as a standard to compare the mechanisms that Spanish, which is the language that will be studied in this dissertation, uses to mark focus. The reason of this is that most studies are done in this language. However, focus marking mechanisms between Spanish and English are different, and some comparisons could not be entirely made, hence the importance of doing research in Spanish to better understand how these different mechanisms affect focus acquisition. In this sense, in the example shown (1), focus falls on the right-most (and hence, the last) constituent in an utterance in both English and Spanish, which tends to be the prepositional predicative of the sentence (Mora-Bustos, 2010, p. 220; Vidal & Leonetti, 2019). This is the most common pattern of prosodic focus in both languages. In this case, regarding its scope in a sentence, this type of focus is classified as *narrow focus*, as it falls in one element of the phrase. However, when answering to questions such as "What happened?", the whole phrase would be new information and hence, the focus. In this case, it is classified as *broad focus* (Gutiérrez-Bravo, 2008).

English is a language with a fixed word order of SVO (Subject, Verb and Object). This means that English speakers will not create, for example, phrases where the Object or the Verb are presented first (or at least not normally). Spanish, in contrast, is a language with a free word order; this means that any element of the phrase may be placed in different positions, and these positions are usually related to a meaning in a particular communicative context (Gutiérrez-Bravo, 2008).

Usually, word order in Spanish is determined by the semantic roles of the verb arguments. In the case of an unmarked utterance with broad focus (that is, a phrase where all its elements are the focus, as shown before with utterances that answer questions like “What happened?”), word order is determined by the hierarchy Agent > Experimenter > Theme/Patient > Place. In this sense, transitive verbs found in this language, for example, have an Agent and a Patient, and hence an order of SVO, which is usually attributed to Spanish language (Gutiérrez-Bravo, 2008). Because of this, it has been argued that this is the “typical” word order in Spanish.

As explained before, prosodic focus, usually in its informative function, is marked at the end of phrases in both languages, that, according to their typical word order, is the Object of the utterance. However, Spanish speakers also place the word that carries informative focus, or the new information, at the end of the phrases when needed, even when this word is not necessarily the prepositional predicative. This does not happen in English because its word order cannot be modified (Gutiérrez-Bravo, 2008; Vidal & Leonetti, 2019). This is the main difference between these languages regarding informative focus marking.

For example, if someone asks, “Who bought the ring?”, Spanish speakers may answer the following:

(2) *Lo compró PEDRO*.

[PEDRO bought it].

As observed, the new information is now the Subject of the sentence and is placed at the end of the phrase by Spanish speakers.

In addition, aside of this placement, Spanish speakers also pronounce this focused word with a greater stress or emphasis. The following section will explain how this emphasis, or acoustic prominence, is marked prosodically in this language.

1.1.3. Acoustic prominence when marking focus in Spanish

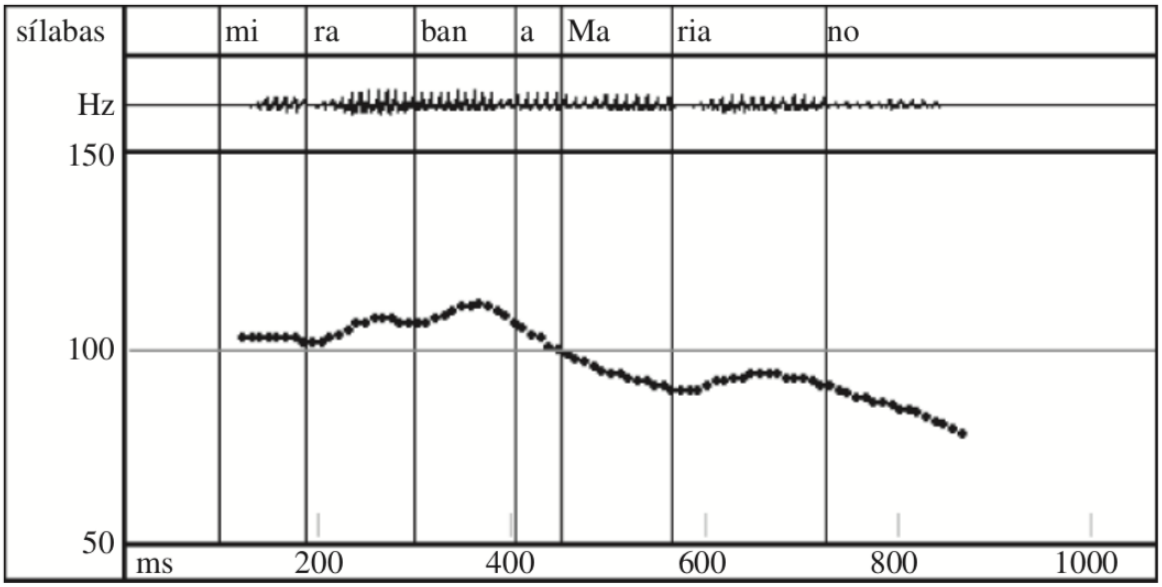
Lexical words (that is, words that have meaning by themselves) in Spanish have a pitch accent in one of their syllables. The accentuated syllable that is more prominent than the others in a spoken phrase is called *nuclear accent*. Nuclear accent is always related to prosodic focus, as it is the last accentuated syllable of the phrase (Gutiérrez-Bravo, 2008).

Accent (and hence, acoustic prominence) is marked through three main elements: syllable duration, intensity, and fundamental frequency (tone) (Face, 2002). *Syllable duration* refers to a longer pronunciation of the vowel that has the nuclear accent and *intensity* means a higher voice volume. *Tone,* in addition, is determined by the frequency in which vocal cords vibrate. When this vibration is high, a high pitch accent is produced; when it is low, a low pitch accent occurs.

The acoustic correlate of the tone is the fundamental frequency of the sound wave (or F0). In turn, F0 is a correlate of word accent. Syllables that bear an accent are the relevant pinpoints of the intonational curve of F0, as such curve raises every time an accented syllable is present. In Spanish, F0, or intonation itself, is used to express pragmatic meanings, and hence is an *Intonational Language*. In the case of focus, the word that represents the new information, for example, has an elevation of the tone on its accented syllable, as shown in Figure 1 (Hualde, 2013):

Figure 1

Example of F0



Note. Example by Hualde (2013). An elevation of the intonational curve or F0 is seen in each accented syllable of the words "*miraban*" (looked at) and "*Mariano*", in the sentence "They looked at Mariano". In this case, "Mariano" is the focused word, and although it has an elevation of the tone in its accented syllable, it is not more prominent than the fist accented syllable of the phrase by its intonational curve, but by its duration.

Tone, duration, and intensity have equal or different relevance to mark the prominence of the nuclear accent (or focus) depending on the type of focus and the language (Face, 2002). In English, the prominence of the nuclear accent is marked with a greater intensity or volume of the voice, as well as an elevation of the tone (Gutiérrez-Bravo, 2008); while in Spanish, even if nuclear accent does have an elevation of the intonational curve of F0 (a higher pitch accent), it is not always more prominent than other accented syllables. Instead, prosodic focus is marked with a greater syllable duration in neutral declarative utterances (as seen in Figure 1) (Hualde, 2013).

But what happens when prosodic focus is used in other contexts that are not informative and does not always appear at the end of the phrase? *Contrastive focus* is an example of this, and it will be explained in the following section.

1.1.4. Prosodic focus in a non-final position in Spanish: The case of contrastive focus

Prosodic focus, as stated, has different pragmatic meanings depending on the context, one of them being informative, where focus falls at the end of the sentence. However, this is not always the case, as demonstrated by one type of focus called *contrastive focus*. Focus is classified as contrastive when, in a conversation, it is used to reject previous information or correct something that was already said (Chen, 2010); in other words, its function is to replace information that the listener has in mind (also named "replacement focus" by Dik, 1997, p. 333).

In this case, acoustic prominence is also used; nevertheless, the element that is focused is not always the last element of the sentence, but the element that needs to be contrasted or corrected, which can be located in any position of the utterance (Mora-Bustos, 2010). For this reason, the meaning of an identical sentence may change depending on the element that is marked with acoustic prominence. This is exemplified in the following utterances from Spanish, as well as their English translations, shown below:

(3) a. *PEDRO compró ese anillo para Juana.*

[PEDRO bought that ring for Juana.]

b. *Pedro COMPRÓ ese anillo para Juana.*

[Pedro BOUGHT that ring for Juana.]

c. *Pedro compró ese anillo PARA JUANA.*

[Pedro bought that ring FOR JUANA.]

In this case, acoustic prominence in “PEDRO” seen in (3a) would indicate that Pedro was the one that bought the ring, and no one else; acoustic prominence in “BOUGHT” in (3b) would indicate that Pedro *bought* (as opposed to, for example, *stole*), the ring; finally, acoustic prominence in “FOR JUANA” seen in (3c) would indicate that the ring was bought for Juana and for no one else. It is important to mention that these examples make sense if we imagine them in a conversation in which the focused elements needed to be emphasized because someone thought that someone else bought the ring (3a), that Pedro stole the ring (3b), or that the ring was bought for someone else that was not Juana (23). Hence, contrastive focus guides the listener’s attention to what the speaker is trying to communicate or correct.

In this line, as explained before, English and Spanish speakers mark prosodic focus using acoustic prominence in the right-most constituent of the utterance (which is the last element of the sentence) (Cinque, 1993). This is different when contrastive focus is used, as seen in (3), where the focused element may be at the beginning (3a), the middle (3b), or the end of an utterance (3c).

Regarding how acoustic prominence is marked in contrastive focus in Spanish, it has been reported that it is more common that F0 raises during the tonic or pretonic syllable and reaches its maximum height inside this tonic syllable (Face, 2002). It is followed by an abrupt tonal drop in the posttonic syllable and, in some cases, by a pause between the focused constituent and the rest of the utterance (Hualde, 2013). Because of this, it is stablished that contrastive focus has a greater acoustic prominence than informative focus (Vidal & Leonetti, 2019).

In addition, the position of the tonal peak, inside the tonic syllable, is relevant to identify the contrasted element, as it is related to emphasis. It is also reported that it is marked with a greater duration of the accented syllable, while intensity (measured with Hz) does not seem to be relevant (Vanrell, Stella, Gili Fivela, & Prieto, 2013). However, in Spanish-speakers from Mexico City, it has been reported that contrastive focus is marked with a greater intensity and duration of the nuclear accent, which are more relevant to identify emphasis than a more prominent F0 (De-la-Mota, Martín-Butragueño, & Prieto, 2010).

English also uses acoustic prominence to mark contrastive focus; however, it is its only resource. In contrast, Spanish speakers may also use an additional mechanism: syntactic movement (or changes in word order inside of an utterance). This may or may not be used, as tonal prominence is always present and guides the interpretation of the utterance (Mora-Bustos, 2010, p. 220). See the following examples:

(4) a. *El anillo que compró Pedro es para Juana.*

[The ring that Pedro bought is for Juana].

b. *PARA JUANA es el anillo que compró Pedro.*

[*FOR JUANA is the ring that Pedro bought].

c. *Pedro fue quien lo hizo.*

[Pedro was the person who did it]

d. *Quien lo hizo, fue PEDRO.*

[*Who did it, was PEDRO.][2]

[2] Note that, in this case, sentences in English with syntactic movement are not possible. An asterisk is used to mark their ungrammaticality.

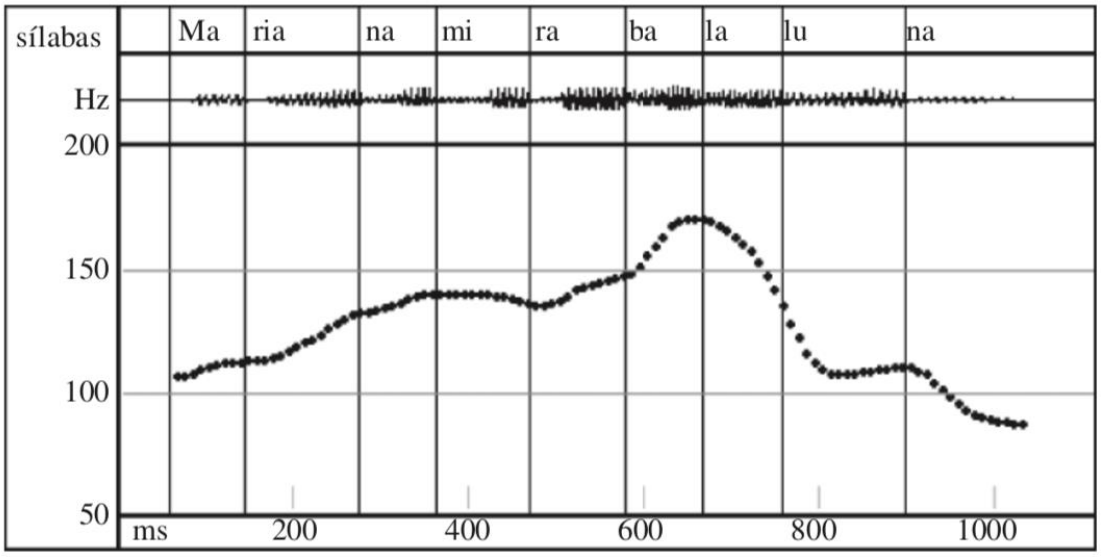
As shown in (3), Spanish speakers may change the word order of an utterance like (4a) or (4c) to place the focused element at the beginning (4b) or at the end of the utterance (4d) in order to emphasise these elements when contrastive focus is used. Acoustic prominence is present as well. In these cases, in an imaginary conversation, it was necessary to point out that the ring was not for someone else, but for Juana (3b) and that the person who did something terrible (we can imagine) is Pedro (3d).

It is important to point out that, in Spanish, contrastive focus is normally marked placing the contrasted element at the beginning of the utterances (Vidal & Leonetti, 2019). This is the most common pattern of contrastive focus for adult speakers in this language. Nevertheless, it can be argued that it is still at the end of the spoken phrase, as Spanish speakers end the phrase with a higher tone or a pause between the focused element and the rest of the sentence in some cases (Dorta-Luis, 2008).

Contrastive focus is one example of prosodic focus being marked at another position of the utterance that is not the end: for example, at the beginning. However, emphasis on a word in this position may also be because of topicalization; that is, because it is information that was mentioned before in the conversation and hence is the theme of what is being said. In these cases, there is a rise of the intonational curve that reaches its peak in the posttonic syllable, as seen in Figure 2 (Hualde, 2013):

Figure 2

Example of Topicalization



Note. This example from Hualde (2013) “Mariana looked at the moon”, answering to the question “What was Mariana looking at?”. Here, “Mariana looked” is the topic and hence, the known information, and “at the moon” the focus or new information answered by the question.

In this case, the intonational curve rises in *miraba* (looked at) because it is the end of the known information (that Mariana was looking at something), and its peak is reached at the end of this last syllable, which, perceptually, may sound as a pause between this word and the rest of the phrase as there is an abrupt fall in the intonational curve after this element.

Nonetheless, it has not been reported how children may interpret focus in different positions in Spanish. As past research has centred in languages like English, German, and Danish, it is critical to determine how prosodic focus acquisition works in languages like Spanish that also use word order to mark contrastive focus. This will help to understand how the acquisition of this aspect of prosody works through the comparison of Spanish with other languages.

In consequence, the present dissertation investigates prosodic focus development in Spanish speakers, focusing on the comprehension of prosodic focus in different positions of the sentence. The investigation is centred around the acquisition of its comprehension, a topic that is still under research. Researching on the acquisition of prosodic focus comprehension is relevant because, in the natural process of language development, language is first understood and then produced. This is the most common pattern: from the acquisition of vocabulary to the acquisition of grammar (Hendriks & Koster, 2010, p. 1887). For this reason, to understand how an aspect of language is acquired, it is pertinent to study the processes involved in its comprehension first.

The next section of this chapter will explore the way in which language acquisition is normally studied, as well as defining comprehension as it will be studied in the present investigation. Two cognitive processes behind prosody comprehension, such as *active verbal memory* and *inhibitory control*, will also be explained in a following section.

1.2. Acquisition of Focus: Comprehension

Language is not a learned ability. Children listen to a limited number of expressions and yet can create an infinite number of novel and creative utterances. Despite individual differences, infants tend to experience the same stages of first language acquisition in the same orderly fashion (Rains, 2004, p. 126). Furthermore, when they are born, babies are capable of distinguishing subtle differences between speech sounds of any language, specialising on their own and loosing this ability at about 6 months of age (Kuhl et al., 2008). The brain's previous preparation to develop a language in the uterus seems to underlie this phenomenon (Rains, 2004, p. 127).

Considering this, language acquisition is studied from two perspectives: comprehension and production (Reyzábal & Santiuste, 2005, p. 129). According to Reyzábal and Santiuste (2005, p.

132), comprehension can be defined as a process that has three relevant moments. The first one is *perception*, in which the information is obtained, and the individual acknowledges it but does not understand it yet (for example, when foetuses and new-born babies distinguish changes in prosody but do not know their meaning yet). The next moment is *analysis*, in which the information, after being perceived multiple times, is linked to a meaning depending on the context (in this stage, for example, children know the meaning of changes in intonation in their parents' voices which may reflect happiness or anger). Finally, the last moment is *use*, in which that information is put into practice (for example, after being able to understand its meaning, children use intonation to express themselves; in this case, language is produced after it is comprehended).

Taking into account these moments of comprehension, this study aims to asses if Spanish speaking children not only *perceive* the changes of intonation present in prosodic focus, but also *analyse* these changes (that is, if Spanish speaking children not only distinguish the tonal prominence in prosodic focus, but also know its meaning and use it as a cue to interpret utterances). To study this, it is precise to define the cognitive processes behind prosodic focus comprehension, which will be explained in the following chapter.

1.2. Cognitive Processes Behind Prosodic Focus Comprehension in Speech

Two cognitive processes involved in prosodic focus comprehension are active verbal memory and inhibitory control. These processes will be explained below.

Memory can be defined as the recall of past information implicitly or explicitly. It can be divided depending on the duration of the retained information in *long-term memory* (for the information that is implicitly or explicitly recalled when some time has passed after it is

obtained) and *short-term memory* (for the information that is usually retained only for seconds or minutes) (Brickman & Stern, 2009, p. 175).

However, the information that stays for a short amount of time in memory can be transferred to long-term memory; not all of it disappears. For this reason, some prefer the term *immediate memory* instead of *short-term memory (*Paller, 2009). One type of information that is kept in immediate memory is verbal information. To understand what we hear when someone speaks, we must keep in memory the words for a short amount of time, grasping the general ideas they communicate. This type of memory can be classified as *active verbal memory (*Neisser, 2014, p. 208).

Usually, when we hear someone speak, general ideas are remembered, and not each utterance with every word that contains it. Neisser (2014, p. 209) explains that people tend to remember chunks of information and no single units as a strategy to retain more information. Everything that is heard seems to follow this process, called *grouping*. Information, words or digits, are reformulated; that means that the subject adds rhythmic clusters or gaps into any spoken list of elements, forming groups of elements that the subject considers are related to each other to remember a whole group instead of each element separately. As a result, information is converted from *echoic* (repetition) to *linguistic* (groups of elements related to each other, forming patterns). These grouping mechanisms used to remember information, according to Neisser (2014, p. 210), could be the same as those involved in speech perception and attention.

This also applies to focus comprehension. In this case, acoustic prominence makes certain elements of the discourse more salient than others, guiding the interpretation of the hearer to what is relevant instead of making them pay attention to every element of the spoken phrases. This helps the hearer to group the relevant elements (as intended by the speaker) of the discourse

and understand the general message. This is further explained by the tendency of human cognition to constantly select relevant elements of speech to make communication more efficient and to process information in a more productive way. In this sense, acoustic prominence in focus is interpreted as a relevant cue by the hearers as the greater articulatory effort in its production differs from the least articulatory effort in the other elements of a spoken phrase. Thus, it becomes relevant information and is maintained in memory (Wilson & Wharton, 2006).

Furthermore, Neisser (2014, pp. 221-223) also argues that rhythm, which is part of intonation, is a crucial element in active verbal memory, thanks to which groups of information are formed. This occurs because rhythms are repetitions and these repetitions create patterns. Remembering a rhythmic pattern does not take room in memory span, it rather creates more room in active memory as the subject memorizes these patterns as a whole and not as single units alone. For example, we do not remember every word someone speaks as a list of disconnected elements, but we remember the central ideas of the utterances we were exposed to. These central ideas are understood thanks to the grouping of words into coherent sentences, where words are related to each other through intonation in speech. As explained before, this would mean that acoustic prominence present in informative focus, which forms part of this rhythmic pattern in speech, helps subjects to identify what part of the utterance is relevant and must be remembered more than others, providing an extra cue to decipher what the speaker is trying to communicate. This cue to identify the relevant part of the utterance occurs thanks to a greater articulatory effort by speakers (Wilson & Wharton, 2006).

Past research (Höhle et al., 2016; Ito, 2018; Szendröi, 2004) has considered working memory (the temporary storage of information in order to handle and process this information; Baddeley, 1992), as the cognitive process involved behind informative focus comprehension. However,

working memory involves remembering certain information to perform a task or to solve problems (Conway et al., 2008). Active verbal memory, in the other hand, does not involve the performance of another task while information is maintained in memory (Baddeley, 2012); it just maintains verbal information in memory, through groups of words or ideas formed by prosody, in order to understand speech (Neisser, 2014). For this reason, the present investigation will make the distinction of active verbal memory being the cognitive process linked to prosodic focus comprehension, because the participants are not performing another task while remembering the central ideas of the utterances they listen to (not always, at least), in contrast with working memory. Working memory may be involved, however, in the process of solving some experimental tasks implemented in past research, but not in comprehension itself.

Active verbal memory explains how language is understood as a whole, including focus, which plays a key role because it is part of the rhythmic patterns any spoken utterance has and emphasises on the relevant part of the discourse. However, another important process is involved in prosodic focus comprehension specifically: *inhibitory control*.

Inhibitory control is the capacity to stop or delay dominant or habitual behaviours that may not be adequate in a given context in order to select another behaviour that is adequate in such given context (Anderson & Weaver, 2009). There are two main types of inhibitory control: *interference control*, the ability to prevent distracting information to intrude and divert the attention of a particular behaviour, and *response inhibition*, the ability to supress an automatic response (Liu et al., 2015). Interference control is related to attentional processes (although it is also relevant to motor processes), while response inhibition is related only to motor processes (Nigg, 2000).

Regarding perception, interference control helps to pay attention selectively to what we choose and override attention to other stimuli (Diamond, 2013). This process has been linked to some aspects of the development of sentence comprehension, although the evidence is scarce (Höhle et al., 2016; Minai et al., 2012; Novick et al., 2005).

Novick and colleagues (2005) have argued that there is a relation between cognitive control (a process that includes control inhibition) and the comprehension of complex sentences, especially when such sentences require a reanalysis of their meaning. The authors used garden-path sentences as an example (ambiguous sentences such as "Put the apple on the towel into the box"). When sentences such as this are being processed, the presence of certain words triggers an initial analysis in which other words are expected to be heard next (in the example above, the presence of "put" at the beginning of the sentence generates the expectation that something will be placed somewhere, and when the words "apple on the towel" are first heard, these are the first elements that fit this expectation). However, the following elements that are heard trigger an inhibition of the first interpretation, which was the most probable, to find an alternative and more possible interpretation in the given context (in the example, "into the box" is then interpreted as the place where the apple is put, and "on the towel" is reinterpreted as the position of such apple).

To have this reinterpretation, cognitive control is needed. Höhle and colleagues (2016) argued that, specifically, interference control handles this process. Following the arguments given by Novick et al. (2005), they propose that, when a dominant interpretation of a sentence is present, interference control is involved in understanding an alternative. This is the case of sentences where the subject becomes the focus in a given context, as the subject is normally the topic and the prepositional predicative, the focus. In this sense, focus produces a modification in the shape

of information the listener has in his memory, guiding its comprehension to what is relevant and away from other possibilities (Chafe, 1974).

In this sense, prosodic tends to fall on the last element of a sentence (the prepositional predicative), which is a dominant or habitual behaviour in languages like Spanish and English. Furthermore, children as young as 2 years of age usually understand tonal prominence as informative focus at the end of sentences (new information), which is a cue they use to learn new words (Grassmann & Tomasello 2007; 2010). Consequently, it is possible to assume this is the standard interpretation children give to tonal prominence at first (as the new information of the discourse, usually at the end of the sentences).

Nevertheless, when prosodic focus in another element of the discourse is used, the interpretation of tonal prominence in the last element of the sentence must be replaced by tonal prominence in another emphasised element. In addition, in Spanish, the usual sentence structure must be inhibited to select an order of words that places the focused element in a relevant position (Mora-Bustos, 2010, p. 220). Consider the examples presented in the first section of this chapter:

(3) a. *PEDRO compró ese anillo para Juana.*

[PEDRO bought that ring for Juana.]

b. Pedro COMPRÓ ese anillo para Juana].

[Pedro BOUGHT that ring for Juana].

c. *Pedro compró ese anillo PARA JUANA.*

[Pedro bought that ring FOR JUANA].

In this case, focus may fall at the beginning of a sentence (3a), at the middle (3b) or at a standard position, the last element of the sentence (3c). Furthermore, Spanish may change the sentence order to place the focused element at the beginning (4b) or at the end of a sentence (4d), changing a regular sentence structure (seen in 4a and 4c), as presented in the following examples:

(4) a. El anillo que compró Pedro es para Juana.

b. *PARA JUANA es el anillo que compró Pedro.*

[Pedro bought that ring FOR JUANA].

c. *Pedro fue quien lo hizo.*

d. *Quien lo hizo, fue PEDRO.*

[PEDRO was the person who did it]

Hence, hearers expect tonal prominence to fall on the last element of the sentence and must inhibit this interpretation to understand a sentence where tonal prominence is used in another element of the utterance (3a, 3b) or an element in a different order (4a, 4b).

For this reason, tonal prominence used alone without syntactic movement at the beginning of a phrase would be even harder for children (even if it is also used by adults). It is also possible that they still do not have this interpretation of prosodic focus at the beginning of a phrase in any circumstance.

Novick and colleagues (2005) also argue that children are guided by the most dominant interpretation of garden-path sentences, at least at 5 years of age, because of cortical immaturity, which is related to cognitive control. Inhibitory control develops through pre-school years (Best

& Miller, 2010): At 4 years of age, children start to show certain success on tasks that measure inhibitory control, and by 5 to 8 years of age, and even until adolescence, there is an improvement in accuracy and efficacy on their performance of such tasks.

For this reason, it would be relevant for this investigation to assess whether there is an interaction between children's development of inhibitory control and their understanding of prosodic focus in another element of the sentence that is not at the end as well; especially considering that, as it will be explained on Chapter 1.3., differences in the development of inhibitory control in pre-schoolers may be one of the problems that affected the interpretation of results in past research. Another problem is the relevance of prosodic focus as a cue to solve the experimental tasks.

1.3. The Assessment of the Comprehension of Informative focus in Language Acquisition

This chapter will focus on one main problem that past research has faced when investigating the acquisition of informative focus: the difficulty to assess its comprehension in children that are below 6 years of age. The first section will explain how this problem has led some researchers to conclude that toddlers use prosodic focus as soon as they produce two-word utterances, at about 2 to 3 years of age, but are not capable of understanding prosodic focus' function until they are 5 years old or more (Cruttenden, 1985; Cutler & Swinney, 1987; Furrow, 1984; Hornby, 1971; Hornby & Hass, 1970; MacWhinney & Bates, 1978; Wieman, 1976). The second section of this chapter will review how other investigators have searched for a way to assess prosodic focus comprehension with experimental designs where prosody is a relevant cue to solve the experiments (Gualmini et al., 2003; Paterson et al., 2003; Szendröi, 2004; Szendröi et al., 2018). Finally, a third section will review how other researchers have focused on investigating the process behind the election of a response in experimental tasks done by

children, using methods like eye-tracking or by assessing Response Times (Chen, 2010; Höhle et al., 2016).

1.3.1. Inconsistencies on the age of acquisition: Problems with the experimental designs

Previous studies not only consider that children produce focus at about 2 to 3 years of age and are unable to comprehend it, but also do not seem to agree on the age in which prosodic focus comprehension is acquired in English-speaking children. These studies range from children being unable to understand focus until they are about 6 years-old (Cutler & Swinney, 1998; McDaniel & Maxfield, 1992; Solan, 1980), to children being unable to understand focus at 8 or even 10 years of age (Cruttenden, 1985; Hornby, 1971). The reason of this variety of results seems to be problems with the experimental designs, as prosody is not a relevant cue to solve the tasks used in the experiments. Hence, prosody comprehension is not actually assessed, as it will be explained below.

Initially, there were studies that assessed prosodic focus comprehension using props in their experimental designs that increase the task's difficulty. In these studies, children were either instructed to recreate sentences using toys (for example, "The camel hit the lion and then HE/he hit the elephant" to assess who performed the action, where prosodic focus on "HE" indicates that the lion hit the elephant rather than the camel) (Solan, 1980), or were expected to grab a particular object whose characteristic was marked with informative focus (for example, "Bert doesn't want to eat the BIG strawberry. What do you think he wants to eat?", where a small strawberry was expected to be grabbed) (McDaniel & Maxfield, 1992). These tasks required extra cognitive processes from children according to Chen (2010): the logical thinking behind the identification of pronouns to subjects (in the case of Solan, 1980), as well as the process involved in identifying each prop after hearing the utterances. This could explain why younger children (below 5 years of age) had problems solving the experiments. In addition, according to

Ito (2018), younger children may interpret differently what experimenters expect from them during the task, especially when toys are involved (for example, thinking that they were taking turns to perform an action with the toy, independently from prosodic comprehension).

Other studies used Picture-Matching Tasks, in which participants had to name which image represented better a series of utterances. This is problematic as well; tasks where children must identify the picture that best represents a sentence may not actually assess prosodic comprehension because children rely on the literal meaning of the sentences they hear as a stronger cue, ignoring prosody (Chen, 2010). Not only that, but they tend to follow a standard interpretation, where focus lies on the last element of the utterance. This is shown in Hornby's (1971) and Cruttenden's (1985) studies. Both investigators argued that even at 8 or even 10 years of age, children are still incapable of understanding prosodic focus.

As an example, Cruttenden (1985) used a task in which children heard sentences like "John's got FOUR oranges" or "John's got four ORANGES". Two pictures were shown, one in which John had four oranges and John's companion, Jane, had two oranges, and other in which John had four oranges and Jane had four bananas. The first sentence matches the first picture because the number of oranges is being contrasted, and the second sentence matches the second picture because the fruits are being contrasted. Children interpreted the pictures like the second sentence, as informative focus usually falls on the object of the sentence. It can be argued that interference control played a role because it was hard to inhibit an interpretation where prosodic focus fell in another element; however, it may also be the case that prosody is irrelevant to solve the task: all the information children needed to answer was in the meaning of the sentences. If this is the case, changes in prosody could be easily ignored, and hence a standard interpretation of focus at the end of a sentence remains as the only option.

The claim that children rely on the literal meaning of an utterance in these types of tasks, and do not use prosody as a relevant cue to interpret what they hear, is further proved in a study from Cutler and Swinney (1987); in this case, in a task where children have to express if a sentence is correct or not. The researchers assessed linguistic processing measuring Reaction Times (RT). They tested children from 3;0 to 8;0 years of age. Two relevant experiments of their investigation are those that assessed prosodic focus comprehension in grammatical sentences (for example, "The nurse brought a CLEAN towel and took away the DIRTY one") and ungrammatical sentences without prosody except for the stressed focus word (for example, "The took and a CLEAN brought the nurse one DIRTY away towel"). Children were expected to press a button when a target word, always focused, was presented (in the examples, "DIRTY" and "CLEAN"). Results showed that all age groups were successful identifying focus when the sentences were ungrammatical (with similar mean response times across groups), but only children below 5 years-old had longer RT when the sentences were grammatically constructed (with mean response times by children younger than 5 years-old at 675 msec versus mean response times by children older than 5 years-old at 375 msec).

The fact that the children were able to identify prosodic focus at all ages (from 3;0 to 8;0 years of age) when the sentences had no meaning, but younger children had a greater difficulty doing so in the case of ungrammatical sentences, suggests that when children younger than 5 years-old must express if a sentence is correct or not, the meaning of the sentence is more relevant than changes in prosody; in contrast, when there is no meaning, they seem to be able to identify and understand the changes in focus intonation.

Cutler and Swinney's (1987) results contrast with other studies that tested prosody comprehension in sentences with and without prosody. These investigations stated that children

see no difference in both conditions, and in consequence, were not able to understand prosody (Lahey, 1974). Nevertheless, the results of these studies could be explained in the same fashion: Children search for the most reliable cue (which is the meaning of the sentence, ignoring its prosody) to answer to tasks in which they must tell if a sentence is correct or not. In these cases, changes in prosody are not relevant for sentence interpretation, and children see no difference if it is present or not.

In conclusion, the discrepancy in age of acquisition reported in the previous investigations could be explained by their experimental design, based on the identification of a sentence with an image or object, as well as tasks in which children must state if a sentence is correct or incorrect. This leads the participants to answer without paying attention to the prosody, considering the literal meaning of the sentences as a more reliable cue to answer. This statement seems to be confirmed in the study by Cutler and Swinney (1987), in which children performed better when the sentences were ungrammatical: if there was no meaning, prosody became relevant. This led the participants to focus on intonation. In addition, when props are used, children find difficult to identify each prop with the utterance and get distracted by the toys, doing something else than what the researchers expected (Chen, 2010; Ito, 2018).

For the reasons stated above, other studies, presented below, tried to solve this problem searching for an experimental design in which prosody was relevant to interpret the sentences used on the experimental tasks.

1.3.2. Prosody as a relevant cue to solve the experimental tasks

Paterson et al. (2006) pointed out that it was necessary to create sentences where the participants had to pay special attention to prosody, especially to intonation in prosodic focus. This is the case of sentences with adverbs such as “only.” This adverb is interpreted as making a

contrast between the focused item and its possible alternatives. However, when placed after the subject, the focus can be interpreted as being on the verb or on the sentence's direct or indirect object, especially if it is read without any prosodic information available. This makes the sentence ambiguous. This ambiguity is solved when there is prosodic information available, as "only" will always be linked to the stressed word of an utterance (the one with tonal prominence) (Reinhart, 1999). This is exemplified below:

(5) a. I only bought chocolate for the boy.

b. I only BOUGHT chocolate for the boy.

c. I only bought CHOCOLATE for the boy.

d. I only bought chocolate for the BOY.

In this case, (5a) is ambiguous and can have (5b), (5c) and (5d) interpretations. This ambiguity is solved with tonal prominence in one of the words that follow the adverb "only". In the case of (5b), the interpretation is that the only action done was buying (not stealing, for example) chocolate for the boy. The interpretation of (5c), in the other hand, is that the only thing bought was chocolate and nothing else. Finally, the interpretation of (5d) is that the chocolate was bought only for the boy and for no one else. It is important to point out that in languages like English and Spanish, the standard interpretation is (5d), as focus normally falls on the object of the sentence (Mora-Bustos, 2010, p. 220), unless other elements are marked with tonal prominence to express, for example, contrast, as it is the case of (5b) and (5c).

The clear link between the adverb "only" and tonal prominence used to mark focus helps adult speakers to understand this type of sentences. However, several experiments showed that

this element makes focus interpretation more difficult for English and Dutch-speaking children (Gualmini et al., 2003; Paterson et al. 2003; Szendröi, 2004). This difficulty may occur because children need to know the meaning of the quantifier and its function; specifically, to what element of the sentence it refers (Chen, 2010; Paterson et al., 2006; Szendröi, 2004). Not only that but, as stated by Gualmini et al. (2003), children usually follow a standard interpretation of focus, and are unable to think differently unless the semantic meaning of the sentence and the context guides them to other interpretation. This means that interference control plays a role. In the experimental tasks, children were unable to inhibit a standard response where the prosodic focus is always interpreted like (5d): With prosodic focus at the end of the sentence.

If experimental tasks that use sentences with the adverb "only" do not make prosodic focus relevant for children, what other strategies to assess informative focus comprehension can be used? Furthermore, when prosodic focus is relevant to children to solve an experimental task, interference control would still play a role in their comprehension of prosodic focus?

One study sheds light into an experimental design that may make prosodic focus a relevant cue to understand the correct meaning of an utterance. Not only that, but it also investigates how this comprehension works in languages that use other methods to mark focus along with tonal prominence, like syntactic movement in Spanish (Mora-Bustos, 2010, p. 220).

Szendröi et al. (2018) searched for a better condition in which participants had to rely on tonal prominence to identify focus and hence correctly interpret a series of sentences. This condition is when the speaker says incorrect information and needs to be corrected by the participant of the experiment (Féry, 2013) using contrastive focus. As stated before, contrastive focus guides the attention of the subject to what the speaker is trying to communicate, making a correction or contrast of previous information. As shown in past research, children's comprehension of

prosodic focus was measured with Picture-Marching Tasks or tasks where children must say if a sentence is correct or not. However, Szendröi et al. (2018) created a task in which children did not have to choose or point out a correct or incorrect sentence or picture, but rather show their prosodic focus knowledge by first, listening to an incorrect item, which was pronounced with tonal prominence, and then use contrastive focus to point out the correct answer.

In the task, the experimenter told the child that he or she needed help to remember a series of images that the infant had in sight, but the experimenter did not. The experimenter said some information incorrectly, and the child felt that he needed to correct him or her. The errors were stressed subjects or objects, and in consequence, the child's answers had to refer to the stressed subject or object (for example, if the experimenter said "The BIRDIE has the bottle, right?", the infant was expected to correct the subject and say "No, the HEDGEDOG does"; and if the experimenter said "The birdie has the BOTTLE, right?", the child was expected to correct the object and say "No, he has the HAMMER") . Children from 3 to 6 years old were tested: 52 children and 16 adults who were English native speakers, 47 children and 11 adults who were French native speakers, and 57 children and 11 adults who were German native speakers. Results showed that, independently of age and language, children from 3 years of age were sensitive to prosodic focus, and hence understood it the same way than adults.

Another revealing aspect of the results is that French (a language that, just as Spanish, not only uses tonal prominence to mark informative focus but also places the most relevant information at the beginning of the utterance), is also sensible to prosodic cues. In the experiment, French children and adults had correct answers referring to the subject when this subject was marked with tonal prominence. These results suggest that children from languages that not only use prosodic cues to mark focus may still use them to interpret the meaning of the

discourse. However, French children also had more answers referring to the object of the sentence than the other children in this subject condition. The experimenters concluded that these results were expected because French uses syntactic movement to place the focused element, the object, at the beginning of the phrase.

Considering these results seen in French children, it is possible that, if a standard response of prosodic focus involves tonal prominence at the end of an utterance, and French speakers, in addition to tonal prominence, also use syntactic movement to place the focused element at the beginning of in some cases, this might mean that inhibiting this response to mark focus on an element that is not at the end of an utterance, without syntactic movement, would be even harder. This might be the case of French-speaking children that are still developing prosodic focus comprehension in another element that is not the end; and hence, it might also be the case of Spanish speakers who also use these mechanisms.

As presented, Szendröi et al. (2018) created a task that guides the attention of children to the prosody of sentences: Using contrastive focus, children had to pay attention on what element of the utterance had tonal prominence and hence had to be corrected or contrasted. Other studies have considered that this undeveloped inhibitory control in children younger than 6 years old may only interfere in prosodic focus comprehension in experimental tasks where children must select a proper answer, but their comprehension is as competent as adults. Because of this, these studies used another strategy to assess prosodic focus comprehension: RTs (as seen in the experiment by Cutler and Swinney, 1987) or the use of on-line measures (experiments that measure the process of the task in real time, such as eye tracking tasks). These types of experiments test linguistic processing while an utterance is heard, and hence measure in a better way prosodic comprehension. Some examples of past research are presented below.

1.3.3. The role of interference control: Experiments that measure linguistic processing

The following experiments showed that children have a similar processing of prosodic focus than adults, at least in the case of German and Dutch (languages in which the experiments where done, where prosodic focus works similar than English).

Chen (2010) tested RTs in a task where children were asked to express if a sentence was correct. She tested 20 Dutch-speaking children from 4 to 5 years old and 15 adults. In the experiment, children saw a series of images, and heard questions and answers to these questions. They needed to say if these answers were correct or not. All the answers were semantically correct, but some had tonal prominence on the focus (correct accentuation) while others had tonal prominence on other element of the utterance, which was abnormal (for example, if the question was "Who has the ball?" the answer had to be "JOHN has the ball", not "John has THE BALL.") Children still had several incorrect responses. It was shown later that this was because they had different interpretations of some of the words' meanings (for example, they rejected utterances as correct answers because certain discrepancies between the pictures and the intended word, like 'grass' instead of 'plant' or 'Father Christmas' instead of 'dwarf'). However, when the intonation was correctly related to focus, both adults and children had faster RT. In contrast, when the intonation was incorrectly related to focus, participants seemed to have detected an intonation pattern that was unexpected, and RT was hence slower. This also suggests a similar processing and understanding of prosodic focus by children and adults.

Chen's (2010) results are relevant because, despite the number of incorrect answers children had in the experimental task, their RTs showed an understanding of prosodic focus changes, as they seemed to detect an incorrect intonation pattern that made harder the election of the response. This suggests that the problem in assessing prosodic focus may be related to answering

the experimental tasks and not the comprehension of informative focus itself. This statement is further analysed in an experiment by Höhle et al. (2016) using an eye tracker.

Höhle et al. (2016) assessed both correct responses as well as gaze patterns. They used experimental tasks with sentences with the adverb "only" in which, as explained before, focus normally falls on the last element of the sentence (Mora-Bustos, 2010, p. 220), which is usually the object. As seen previously, sentences with the adverb "only" were difficult to interpret by children. Höhle et al. (2016) considered that the reason behind this difficulty lies on their undeveloped inhibitory control when they had to answer the experimental tasks. In the case of prosodic focus in sentences with the adverb "only", this would mean that children will always choose the default placement of focus on the object of the sentence because they are still incapable of inhibiting this answer to choose another one. This theory is tested in two experiments.

In the first experiment, 17 German speaking children with a mean of 4;5 years of age, and 17 adults, were tested. They used a Picture-Matching Task as the ones used before. Pre-subject and pre-object sentences with "only" were used (for example, in the pre-subject condition, children heard: "The mouse, the mole, and the duck have a balloon. Only the ELEPHANT has a kite", which would match to a picture in which the elephant was the only animal to have a kite, but this elephant was also holding other objects; and in the pre-object condition, children heard "The mouse, the mole, and the duck have a balloon. The elephant has only a KITE", which would match to a picture in which other animals have kites, but the elephant was holding a kite and not any other objects). Experimenters also assessed the gaze patterns of the participants. It was expected that the gaze of the participants focused on the characters or on the objects these characters were holding depending on the condition, after they heard the sentences and were

analysing the images. Hence, the areas of interest were the characters in the pictures and the objects they carried.

Results showed that children had a great amount of "yes" (as in, "the sentence matches the picture") responses, which could be explained by chance. However, their gaze patterns where the same as adults: focusing on the characters in the pre-subject condition and on the objects on the pre-object condition. These results suggest that their understanding of prosodic focus is similar to adults, but other factors intervene on their final answer.

For this reason, the experiment was replicated a second time, with different participants of the same age (this time, 34 children), adding a test of inhibitory control (specifically, of interference control) in a task in which children had to press a left or a right button depending on the side a fish on a screen was facing (which is known as "Flanker Task"). This fish had two fishes at each side that would be facing the same or a different side as the target fish. The results of this experiment regarding responses and gaze patterns were identical to those of the first experiment; nevertheless, difficulties in interpreting pre-subject "only" were related to low scores on the inhibitory control experiment. This could explain the results shown on past literature: Children have the same linguistic processing than adults, but their poor inhibitory control prevents them from selecting another answer that is not the standard, where focus is always placed on the object (or the last element) of the sentence. Even if they detect and understand the meaning of a change of intonation to a new constituent of the utterance, they are unable to inhibit a standard answer (with focus on the last element of the utterance) in this type of tasks.

The results seen in Höhle et al. (2016) are interesting because they support the hypothesis that an undeveloped inhibitory control may interfere with the results of the experiments evaluated so far, not only with utterances that use the adverb "only", but also with those that assessed

prosodic focus comprehension where focus may fall in another element of the utterance that is not the last one.

1.4. Rationale & Problem Statement

The present investigation therefore aims to contribute to the research about pre-schoolers' development of the comprehension of prosodic cues in focus, by assessing such development in Mexican Spanish-speakers.

As stated, prosodic focus marks, through tonal prominence (a greater duration, intensity or high pitch of the accented syllable of the focused word), the information of a spoken phrase that the speaker wants the hearer to pay attention to (Dorta-Luis, 2008); sometimes when this information is new to the hearer (Hualde, 2013; Ladd, 2008; Vallduví & Engdahl, 1996). This information, because of its acoustic properties, tends to be the last element of such phrase, usually the object (Gutiérrez-Bravo, 2008; Vidal & Leonetti, 2019; Mora-Bustos, 2010).

This is the most common pattern of intonation in most declarative spoken phrases in Spanish (Hualde, 2013; Vidal & Leonetti, 2019; Mora-Bustos, 2010). Spanish-speakers even tend to move the relevant information at the end of the spoken sentence when needed (Gutiérrez-Bravo, 2008; Vidal & Leonetti, 2019); and even children, from 2 years of age, have this interpretation of tonal prominence at the end of sentences, related to novel words (Grassman & Tomasello, 2007; 2010), as adults make a greater articulatory effort to signal the information that the child must attend to in a conversation in child-directed speech, usually with an informative meaning (Villalobos-Pedroza, 2019).

However, some focused elements may be placed in any position of a phrase. Contrastive focus in Spanish, for example, is usually placed at the beginning or at the end of sentences, and it is marked with a greater intensity and duration of the accented syllable of the focused word, as

well as a more abrupt raise of the tone in such syllable (De-la-Mota, Martín-Butragueño, & Prieto, 2010; Hualde, 2013). And emphasis in other elements, such as the beginning of the phrase, may also be interpreted as topicalization, to cite another example (Hualde, 2013).

Furthermore, as defined before, control inhibition is the cognitive ability to stop or delay a dominant behaviour, or to override attention to a dominant stimulus, to select another behaviour or pay attention to another stimulus that is more appropriate in a particular context (Anderson & Weaver, 2009; Diamond, 2013). This is specially the case of interference control, a type of inhibitory control that is related to cognitive processes (Nigg, 2000) – such as selecting focus on an element that is not, for example, the last element of the phrase, which is its usual placement in most declarative phrases.

It was also stated that interference control was linked to sentence comprehension, as problems in the comprehension of phrases that require a reinterpretation are related to an inhibitory control under development in pre-school children (Höhle, 2016; Minai et al., 2012; Novick at al., 2015). This is especially true since inhibitory control is still developing through pre-school years, as children start to show some success solving experimental tasks that assess inhibition by 4 years of age (Best & Miller, 2010).

As presented in the literature review, Höhle et al. (2016) showed that a poor inhibitory control was related to 4-year-old children's performance in experimental tasks that assessed prosodic focus comprehension in sentences with the adverb "only". This adverb is usually related to the object of a sentence. Pre-schoolers who performed worse in the inhibitory control tasks also tended to produce wrong answers when "only" was placed before the subject of the sentence, as they still interpreted focus on the object rather than on the subject. Pre-schoolers were unable to inhibit this dominant and standard verbal response, even when their gaze patterns showed an

understanding of the changes in intonation. Taken these results into consideration, it is plausible that this will also occur in other types of sentences where prosodic focus is presented without "only".

Examples of sentences without "only" are present in recent studies that were able to assess prosodic focus comprehension in pre-schoolers between 3 to 6 years of age of languages like English, German, French and Dutch (Chen, 2010; Höhle et al., 2016; Szendröi et al., 2018), where French children of those ages still had a harder time selecting an answer where prosodic focus was not at the end (or was the main object) of an utterance.

French uses not only tonal prominence, but also syntactic movement to mark prosodic focus, placing it, for example, at the beginning of spoken phrases (for example, "*Non, c'est SA MÈRE qu'il a appelée*" ["No, the one he called was HIS MOTHER"], Authier & Haegeman, 2017). This is similar to Spanish, that also uses syntactic movement to mark contrastive focus at the beginning of spoken phrases (for example, "*DE JUANA es el anillo*" ["The ring belongs TO JUANA"]) (Gutiérrez-Bravo, 2008; Vidal & Leonetti, 2019; Mora-Bustos, 2010). Both languages share this similarity in contrastive focus marking.

In consequence, if, for adults, prosodic focus in French and Spanish is more commonly placed at the beginning of a phrase through syntactic movement (Authier & Haegeman, 2017; Gutiérrez-Bravo, 2008; Vidal & Leonetti, 2019; Mora-Bustos, 2010; Szendröi et al., 2018), it is possible, then, that pre-schoolers from these languages have a harder time inhibiting a standard interpretation of focus at the end of phrases to understand prosodic focus at the beginning without syntactic movement (only with intonation). Furthermore, pre-schoolers would listen more frequently to prosodic focus, in general, at the end of spoken phrases (Grassman & Tomasello, 2007; 2010).

Taking this information together, the present investigation assesses prosodic focus comprehension in different positions of an utterance in 4 and 5-year-old Spanish-speaking pre-schoolers, considering that at this age inhibitory control is still developing but children start to show some success in the experimental tasks (Best & Miller, 2010). Differences in comprehension of prosodic focus might be related to this fact (Höhle at al., 2016), especially in languages like Spanish, in which prosodic focus comprehension in positions of the sentence other than the end might be even harder for the reasons presented before. In addition, Spanish speaking pre-schoolers have consolidated their syntax development by this age and continue to do so until they are 5 years old (Ojeda, 2000), which means that this consolidation will not intervene in their understanding of the sentences presented. Furthermore, it is still unclear if pre-schoolers that speak languages like French of Spanish have a competent understanding of prosodic focus in different positions at that age (Szendröi at al., 2018).

Following this line of argument, will Spanish speaking pre-schoolers be able to use tonal prominence as a cue to understand declarative sentences that have prosodic focus at the beginning and at the end of the utterance from 4 to 5 years of age? If there are individual differences in performance, will there be a relation between interference control and the way they give a standard response, placing focus at the end of an utterance, across conditions in the experimental tasks?

Experiment 1 introduces a task in which prosodic focus is relevant, similar to the one used by Szendröi et al. (2018) but adapted to Mexican Spanish speakers, where the participant were encouraged to use contrastive focus to correct the experimenter and children had to pay attention on what element of the utterance had tonal prominence (the beginning or the end of the sentence), and hence had to be corrected. In Experiment 2, the same task was implemented but

adding an eye tracker, which was used to test linguistic processing, since it is a more reliable tool to evaluate prosodic focus comprehension (Höhle et al., 2016). According to Ito (2018), the advantage of this method is its capacity to trace spontaneous reactions to utterances the participants listen to while they see an image on the screen, even before they have to react to the command of an experimental task. The timing of eye fixations on an image on a screen as participants are listening to a sentence reveals their sensibility to prosodic cues. Gaze direction, in the other hand, can determine if the tonal prominence used in contrastive focus is correctly interpreted. This technique, hence, is useful to assess prosodic comprehension in real time as the participant is listening to an utterance.

In addition, eye-tracking in RT studies (Chen, 2010; Höhle et al., 2016) suggest that pre-school children have a competent comprehension of prosodic focus but have problems selecting a final verbal answer in the experimental tasks in languages like German or Dutch. Because of this, the contrast between both experiments would show if, in Spanish-speaking pre-schoolers, there is a problem in the comprehension of the sentences with prosodic focus presented or only on giving a verbal response in the experimental tasks.

To do so, Experiments 1 and 2 were compared by assessing interference control in both experiments. The task that was used is a “Flanker Task”, as it assesses interference control from a cognitive perspective (Nigg, 2000), has been adapted to pre-schoolers (Rueda et al., 2004) and is used to compare its relation to contrastive focus in pre-school children (Höhle et al., 2016). The comparison between experiments was done, first, by assessing interference control during Experiment 2 (that assesses eye fixations and gaze patterns), to evaluate its relation to comprehension itself, and during Experiment 1 (where pre-schoolers have to use contrastive focus to correct the experimenter) to evaluate its relation only to the final answer on the

experimental tasks. If low scores on the response inhibition task are related to a preference to give a standard answer (where prosodic focus is at the end, or the prepositional predicative of the utterance) in Experiment 1, it is possible that an undeveloped inhibitory control is related to the way pre-schoolers choose their final verbal answer in experimental tasks. In contrast, if low scores on the response inhibition task are also related to a gaze preference to the objects shown in Experiment 2 (which would represent the last elements of an utterance and hence, a standard response), it is possible then that an inhibitory control still under development is not only related to the way pre-schoolers choose their verbal responses, but to prosodic focus comprehension itself, which might still be under development as well.

Additionally, Spanish speakers from central-western Mexico were studied, considering this is mostly the population that is accessible through the social media accounts and databases of the Laboratory of Psycholinguistics of the National Autonomous University of Mexico (*Universidad Nacional Autónoma de México*, UNAM). In this region, prosodic focus' production in adults and children has been studied before, which is relevant for the creation of the experimental stimuli (De-la-Mota, Martín-Butragueño, & Prieto, 2010; Villalobos-Pedroza, 2021). Furthermore, as stated by Villalobos-Pedroza (2021), it has been reported that preschool children of the centre of Mexico may produce prosodic focus in non-canonical positions from 23 to 28 months of age.

Talking about the implications of the present investigation in language acquisition research in general, studying the development of prosodic focus comprehension is relevant because several studies (De Carvalho et al., 1991; De Carvalho et al., 2018; Garf Estes & Hurley, 2013; Ma et al., 2011) suggest that sensibility to prosodic cues facilitates the acquisition of new vocabulary, as pre-schoolers use informative focus to allocate their attention to what the speaker marks as the new or relevant element of the discourse (Grassmann & Tomasello, 2001;2010; Ito, 2018, p. 252;

Villalobos-Pedroza, 2019), and hence the use of an exaggerated prosody may cause an increase in children's vocabulary acquisition. For this reason, investigating focus comprehension's development would also help to understand the mechanisms behind vocabulary comprehension as a part of the language acquisition process in general.

Otherwise, considering the practical applications of this investigation, studying the acquisition of prosodic focus comprehension would also help to create better learning instruments to develop linguistic abilities such as vocabulary comprehension, but also reading, as focus guides the attention to relevant information of the discourse. According to Baeza-Álvarez and Rodríguez-Maldonado (2011, p. 13), in Mexico, where this investigation was conducted, there are few studies about the development of prosody in preschool children. Besides, investigating the typical development of different prosodic aspects would also contribute to create effective interventions that could help children with language impairments.

Children with a typical language development have high sensibility to prosodic information. Furthermore, prosodic focus is emphasized in child-directed speech (Szendröi et al., 2018; Villalobos-Pedroza, 2019). Regarding children with language impairments, studies show, for example, that children with autism spectrum disorders pay more attention and respond better when there are prosodic changes, such as intonational changes in the voice of the speaker, than when there are not (Lamers & Hall, 2003). In addition, therapies that use prosody as a tool to improve the communication of children with Down Syndrome are currently under development (Adell et al., *n.d.*; Kumin, 2014). In this sense, prosodic focus helps to find the element of the utterance the speaker wants to communicate, which would also help to create better tools to improve the communication of children with and without language pathologies.

2. Methods & Results

2.1. Sample

The present investigation assessed 42 Spanish-speaking pre-schoolers (24 female) of 4 to 5 years of age ($\bar{x}$= 4:6, *SD*= 0.50; ranging from 4;0 to 5;11). All participants were Spanish speakers from central-western Mexico, a dialectal zone proposed using phonetic data by Martín-Butragueño (2014), composed of the states of Aguascalientes, Colima, Guanajuato, Guerrero, Hidalgo, Jalisco, Mexico City, Michoacán, Morelos, and Querétaro. In this case, 28 participants were from Mexico City, 7 from Aguascalientes, 4 from the State of Mexico (considered as part of the same Metropolitan zone as Mexico City), 1 from Guerrero and 1 from Hidalgo.

Pre-schoolers were divided into two groups: the Subject Condition (SC), where prosodic focus is placed on the subject of the utterance; and the Prepositional Predicative Condition (PPC), where prosodic focus falls on the prepositional predicative. They were divided randomly into these two conditions (SC, *n*= 20; PPC, *n*= 22). In addition, the same participants took part in Experiment 1 and Experiment 2, which were conducted two weeks apart.

Ideally, pre-schoolers should be recruited using probabilistic cluster sampling, since that is the best way to find children of these ages in Mexico. However, the sample could not be entirely probabilistic, since there are parents who, for various reasons, would or could not accept to participate in the experiment. In addition, the experiments were carried out online due to the COVID-19 pandemic. Thus, volunteers that had access to a computer were sought through the UNAM Laboratory of Psycholinguistics, and through social media. Considering this, the population consisted mostly of pre-schoolers of middle-class parents who had access to a computer with webcam. The sample, hence, is less representative of the population, but it was a viable option for the reasons mentioned above.

Furthermore, per sample size calculations provided by G*Power (Faul et al., 2007), a sample size of 614 participants was required considering a small effect size of 0.1, an alpha value of 0.05 and an actual power of 0.80. This value was calculated considering this is a correlational study with one-tailed hypotheses (stated below). It was difficult to obtain such number of participants, but thanks to online experimentation, a total of 42 participants were finally recruited.

2.1.1. Inclusion criteria

Pre-schoolers were assessed to discard any language problems with respect to language comprehension by assessing their receptive vocabulary according to their age. For this matter, the Spanish-Bilingual Edition of the Receptive One-Word Picture Vocabulary Test-5 (ROWPVT-4: SBE) was used. ROWPVT-4: SBE is a receptive vocabulary test and a verbal aptitude screening that covers a wide range of ages (from 2 to 70+ years of age). It evaluates the ability to match a spoken word, in either English or Spanish, to an image. Depending on the number of words identified, an age rank is calculated (Martin, 2013). All children had to have a standard score of 100 points (which is the midpoint of their age group or a percentile of 50, according to the manual) or at least 15 points above or below that score (+ - 1 *SD*) to participate.

Results from the ROWPVT-4: SBE showed that most pre-schoolers had an age of comprehension similar or above their chronological age ($\bar{x}$=5;4 years, *SD*=0.96), ranging from 3;6 to 7;1 years-old. All but one of them had percentile scores of 37% or above, within 1 SD (Martin, 2013), meeting the inclusion criteria (n = 41).

Furthermore, parents filled a sociodemographic questionnaire, stating their educational level, establishing if the child has any cognitive, vision or hearing impairments, along with other information of the child such as their name, date of birth, gender, and languages spoken to their child. Children could not have any hearing, vision or cognitive impairments to participate in the

experiments, and this was assessed in such questionnaire. None of the 42 pre-schoolers presented any auditory, cognitive or language problems as reported by their parents.

2.1.2. Exclusion criteria

Children whose scores in the ROWPVT were below the standard score of 100 with a SD of 1 and had any cognitive, vision or hearing impairments were excluded from the experiments. In addition, participants who did not have the recommended system requirements as specified by the platform RealEye (as shown in the section "Experimental Scenery" below) were not able to participate in the study.

One child had an age of comprehension (2;7) below their chronological age (4;10), with a percentile score of 11%; because of this, the child was excluded from the study, leaving a total of 41 participants in the final sample.

2.1.3. Elimination criteria

Children who did not participate in both experimental sessions, failed to understand the practice trials, or had answers below chance, were eliminated from final analyses.

2.2. General Objective

To compare the comprehension of prosodic focus in the subject or the prepositional predicative of declarative utterances in Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age through eye gaze and verbal responses, as well as to identify the relation of each type of response with interference control scores.

2.3 Specific Objectives

1. To evaluate prosodic focus comprehension in Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age through eye gaze and verbal responses.

2. To compare eye gaze and verbal responses given by Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age to utterances where prosodic focus is embedded in the subject or in the prepositional predicative of the utterances.
3. To compare differences between eye gaze and verbal responses to the subject of the utterance in Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age.
4. To evaluate interference control in Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age.
5. To identify the relation between lower scores in the interference control task and the errors in the verbal response task and the eye-tracking task to the subject of the utterance in Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age.

2.4. General Hypothesis

Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age will have correct eye gaze responses to prosodic focus in the subject and the prepositional predicative of declarative utterances, but will have a greater number of correct verbal responses to the prepositional predicative than to the subject of these utterances, which will be related to lower interference control scores.

2.5. Specific Hypotheses

- $H0_1$: The SC and the PPC will not have differences on the number of correct verbal responses.
- $H1_1$: The PPC will have a greater number of correct verbal responses than the SC.
- $H0_2$: The SC and the PPC will not have differences on the number of correct eye gaze responses.

- $H1_2$: The PPC will have a greater number of correct eye gaze responses than the SC.
- $H0_3$: The SC will not have differences on the number of correct verbal responses and the number of correct eye gaze responses.
- $H1_3$: The SC will have a greater number of correct eye gaze responses than correct verbal responses.
- $H0_4$: Higher inhibition responses will not be related to correct verbal responses to the SC.
- $H1_4$: Higher inhibition responses will be related to correct verbal responses to the SC.
- $H0_5$: Higher inhibition responses will not be related to correct eye gaze responses to the SC.
- $H1_5$: Higher inhibition responses will be related to correct eye gaze responses to the SC.

2.6. Experimental Design

A cross-sectional study was carried out with two experiments that assessed prosodic focus comprehension in pre-schoolers of 4 and 5 years of age and its relation to interference control. It is a quasi-experimental study because the participants were selected according to certain characteristics (such as age, lack of cognitive delay or audition problems, a receptive vocabulary according to their age, and access to a computer with webcam).

In the first experiment, pre-schoolers were expected to use contrastive focus to correct the prosodic focus used by the experimenter in the subject or the prepositional predicative of an utterance. If pre-schoolers tended to correct the utterance marking prosodic focus on the

prepositional predicative more often, even when the subject needed correction, then they were following a standard response of prosodic focus at the end of the utterance. If this pattern of responses is related to low scores of the interference control task, then it would mean that there is a relation between prosodic focus comprehension and an inhibitory control under development.

However, it is possible that inhibitory control under development is related to the verbal responses pre-schoolers give in such experimental tasks and not comprehension itself. For this reason, Experiment 2 assessed their gaze patterns and eye fixations to examine their prosodic focus comprehension. If there is still a pattern in which their eyes look more at the prepositional predicatives than the subjects, even when the subject is marked with prosodic focus, and this data is related to low scores in the interference control task (a larger number or errors and RTs, as it will be explained in the following sections), this would confirm that there is a relation between poor prosodic focus comprehension and inhibitory control under development. It is, in consequence, a correlational study.

It is important to point out that both experiments, in other conditions, would have been done at the same time as a single experiment, where participants would look at an eye-tracker screen while they give a verbal answer. However, as this investigation took place during a the COVID-19 pandemic, eye-tracking data was gathered online through the participants' webcams. Because of this, and to have reliable data, children must be as still as possible, without talking, during Experiment 2. This is the reason why it was decided to have two separate experiments.

The order of application of the tasks was slightly different across participants. The application consisted of two sessions of 30 minutes each, approximately. In the first session, the ROWPVT was applied first, as it assesses receptive vocabulary comprehension, and children whose age of comprehension was below their chronological age were not invited for the second session of the

experiment. This task also served as a warm-up phase. After the ROWPVT, either Experiment 1 or 2 was applied; the order of application was randomized. In the second session, the remaining Experiment 1 or 2 was applied first to vary the order of presentation from the last session. Then, the Flanker task was applied at the end. Table 1 shows an example of the order in which the tasks and instruments were applied in both sessions:

Table 1

Procedure of Experiment 1 and Experiment 2

Session	Order of application	Instrument or task	Approximate duration	Purpose of application
0	0	Informed consent and sociodemographic questionnaire form	5 minutes	To inform about the experiment and to obtain the child's and parents' general information.
1	1	ROWPVT	20 minutes	To assess if the child's language level is typical for their age.
1	2	Experimental task from Experiment 1 or 2	15 minutes	To assess prosodic focus comprehension.
2	3	Experimental task from Experiment 1 or 2	15 minutes	To assess prosodic focus comprehension.
2	4	Flanker Task	8-10 minutes	To assess interference control.

Note. The informed consent and sociodemographic questionnaire were given to the parents prior the session via email or *WhatsApp*. The rest are instruments or tasks answered by the child during the session. The order of presentation of Experiment 1 and 2 was randomized across participants.

2.6.1. Stimuli of Experiment 1 and 2

The stimuli consisted of pictures of six animate or six inanimate objects presented at the same time, with utterances that refer to these images. The words related to the images were known by 30-month-old Mexican Spanish-speaking infants according to the Norms for the MacArthur–Bates Communicative Development Inventories for Mexican Spanish-speakers (Jackson-Maldonado et al., 2003; Weisleder & Fernald, 2013). These words had at least two syllables to easily highlight prosodic prominence, as changes in the intonational curve are seen with more clarity from one syllable to the other. In addition, such words contained only voiced consonants, as well as accents in the penultimate or antepenultimate syllable to facilitate the analysis and control of the intonational curve (Vanrell et al., 2013), as shown in Table 2. It is relevant to mention that the suitability of phonetical characteristics for intonational analysis was considered more relevant than the percentage of familiarity because prosodic cues, which are the focus of this study, were measured and controlled.

A list of the words used in the audio stimuli is presented below, compiled from the Wordbank (http://wordbank.stanford.edu/publications):

Table 2

Stimuli for Experiment 1 and 2

Category	Word	Phonetical transcription	English translation
Animate objects	*Araña* (.87)	[a.'ɾa.ɲa]	Spider
	Hormiga (.77)	[or.'mi.ɣa]	Ant
	Rana (.74)	['ra.na]	Frog
	Gallina (.73)	[ga.'ʝi.na]	Hen
	Búho (.45)	['bu.o]	Owl
	Ardilla (.40)	[ar.'ði.ʝa]	Squirrel
Inanimate objects	*Mano* (.92)	['mã.no]	Hand
	Huevo (.92)	['we.βo]	Egg
	Uva (.68)	['u.βa]	Grape
	Globo (.84)	['glo.βo]	Balloon
	Llave (.81)	['ʝa.βe]	Key
	Libro (.73)	['li.βɾo]	Book

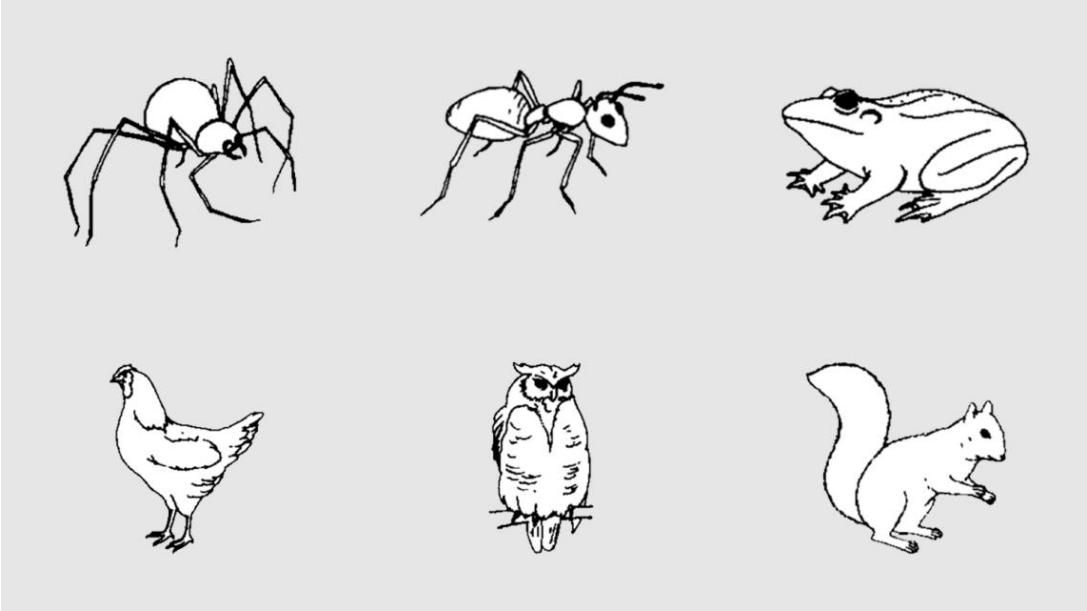
Note. In parenthesis, the proportion of familiarity of Mexican Spanish-speaking infants at 30 months of age. The phonetic transcription corresponds to Mexican Spanish.

The images that represent these words were presented three times in different positions (left, centre or right, either above or below the screen). Animate and inanimate stimuli were presented separately to eliminate the bias in which the topic of a sentence tends to be an animate object and the focus an inanimate object (Szendröi, 2018). This way, the child saw at a given moment only animals or only objects on screen, and the focused element could be any of the animals, or any of the objects presented.

The number of items presented at the same time on screen was six since two pairs of elements, one below the other, were contrasted in the conditions (one element of each pair being the focus and the other one, the topic), and the last pair was the distractor, to ensure the comparison was made between the items (Szendröi, 2018), as seen in Figure 3:

Figure 3

Example of Stimuli of Experiments 1 and 2



Describing the images' characteristics, neither stood out more than the other. To achieve this, clipart vector images in black and white were used. These images had a resolution and size of 440x340 px to maintain symmetry in a background of 1920x1080 px, which is the size of a screen with an optimal resolution of 1080p (this standard was used to transform and compare the data of the different screens that the participants had). Each image was centred in one of six positions. Furthermore, each image appeared a single time in one of the six positions at one given moment on screen, and neither image appeared next to the same image in the same

position. These resulted in a total of 12 combinations taking into consideration both animate and inanimate objects (six combinations each).

The words and the images that represent these words were validated for Mexican Spanish speaking pre-schoolers, as shown in *Appendix A*.

2.6.1.1. Stimuli presentation. Each combination of images was accompanied by three different utterances each; in the case of the experimental items, such utterances contained prosodic focus in the subject or in the prepositional predicative, (for example, “The SPIDER is above the owl, right?” or “The spider is above the OWL, right?” as shown in Table 3). Declarative sentences with the most common word order in Spanish (Subject, Verb and Object) were used. In this type of sentences, prosodic focus may be placed in any element in different contexts in a natural way without further syntactic cues (Gutiérrez-Bravo, 2008). Furthermore, this word order allows comparison to studies in English. In addition, the word order also helped ensure that the distinction between focus on the subject or the prepositional predicative was clear enough on the acoustic analysis of the intonational curve, as these elements are at the opposite sides of the utterance. Examples of the audio stimuli, as well as the expected responses given by pre-schoolers, are presented below in Table 3 and Table 4:

Table 3

Example of Audio Stimuli for Experiment 1 and 2

Subject Condition	Prepositional Predicative Condition
La ARAÑA está arriba del búho, ¿verdad?	*La araña está arriba del BÚHO, ¿verdad?*
[The SPIDER is above the owl, right?]	[The spider is above the OWL, right?]
La HORMIGA está arriba de la ardilla, ¿verdad?	*La hormiga está arriba de la ARDILLA, ¿verdad?*
[The ANT is above the squirrel, right?]	[The ant is above the SQUIRREL, right?]
La RANA está arriba de la gallina, ¿verdad?	*La rana está arriba de la GALLINA, ¿verdad?*
[The FROG is above the hen, right?]	[The frog is above the HEN, right?]

Table 4

Example of Expected Responses for Experiment 1

Subject Condition	Prepositional Predicative Condition
No, está la HORMIGA.	*No, de la GALLINA.*
[No, the ANT is.]	[No, it is (above) the HEN.]
No, está la RANA.	*No, del BÚHO.*
[No, the FROG is.]	[No, it is (above) the OWL.]
No, está la ARAÑA.	*No, de la ARDILLA.*
[No, the SPIDER is.]	[No, it is (above) the SQUIRREL.]

This type of copulative sentences (with the verb "*estar*" or "to be"), in addition to the adverb "*arriba*" or "above" was selected because it was the simpler way to describe the images shown. It made the task easier to understand, as the sentences directly described the images and their position on screen, making them easier to compare for pre-schoolers.

In addition to these test items, control items were used, in which the audio stimuli consisted of correct statements about the images (such as "*La ARAÑA está arriba de la gallina, ¿verdad*?" ["The SPIDER is above the hen, right?"], referring to Figure 3 above). These items tested if the pre-schoolers correctly understood the utterances and the instructions of the task. Furthermore, filler items were also implemented with a neutral intonation; that is, without narrow prosodic focus, or focus on all the utterance (for example, "*El búho está abajo de la araña ¿verdad?*" ["The owl is below the spider, right?", referring to Figure 3 as well]. Half of these statements correctly described the images and half described them incorrectly. The purpose of these items was to limit learning strategies from the participants.

There were 36 experimental trials in total, with 12 experimental items, 12 control items and 12 filler items. In each case, the animated or inanimate object was the subject and prepositional predicative of the utterances two times in either set of items. For a complete list of stimuli, see *Appendix B.*

All utterances were recorded by a female Mexican Spanish-speaker using a Shure MV51 USB condenser microphone at an approximate distance of 30 cm from the speaker, using the program Adobe Audition in stereo at a rate of 44100 Hz in 16 bits. All audio stimuli had 1000 ms of silence at the beginning, so the participants had some time to examine the picture on the screen. Tonal prominence, as well as the duration of the tonic syllables, of all recordings was controlled using Praat (Boersma & Weenink, 2006), as shown in *Appendix C* along with all the recording procedure.

In addition, to adapt the stimuli to Experiment 2, which requires the use of an eye-tracker, the sentences were modified to have the same duration, making sure each section played at the same millisecond. The purpose of this was to enable the analysis of eye-tracker data, knowing where

the participants gaze was at a given moment in time. For this matter, the sentences were divided in article (*El/La*, “The”), subject (animal or object), verbal phrase (*está arriba de la/el* “is above the”), prepositional predicative (animal or object), silence before the final question, the final question *(¿verdad?*, “right?”), and 2000 ms of silence at the end, which is recommended to analyse the gaze after the prepositional predicative of the sentence is heard (Höhle et al., 2016). Figure 4 and Figure 5 show the timestamps of each section; the image appears from the beginning to the end of the utterance:

Figure 4

Timestamps of Sentence Sections of Auditory Stimuli in the Subject Condition

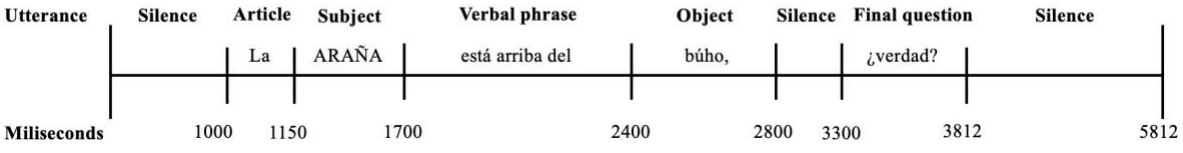
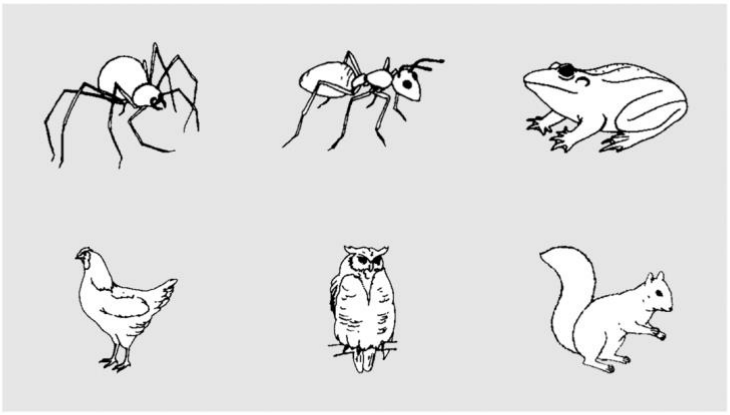
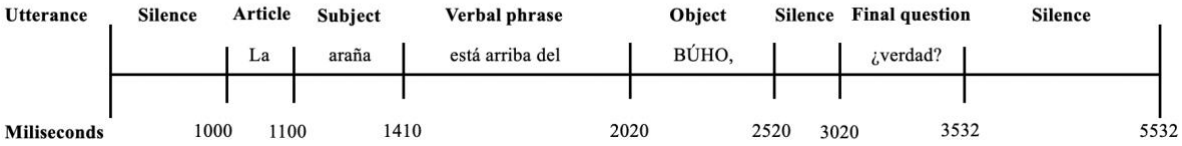


Figure 5

Timestamps of Sentence Sections of Auditory Stimuli in the Prepositional Predicative Condition





The duration of each segment was controlled so that the duration was, at most, 30% slower or faster than its original form, to ensure that the natural quality of the voice would be maintained. Furthermore, a validation (n=5) was made to ensure the stimuli sounded natural, asking the participants if the voice sounded natural or not in each stimulus. All sentences were classified as natural by 70% or more of the participants.

Additionally, another validation of the stimuli was performed with an adult population. The purpose was to make sure the stimuli was understood by adult speakers as well as making sure the prosodic cues were relevant for the interpretation of the experimental tasks. This validation can be consulted in *Appendix D*.

2.7. Experimental Scenery

Because of the current Pandemic, the experimental sessions were planned to be carried out online. In the case of Experiment 1, the test was carried out via videoconference and via the platform "Gorilla" (https://gorilla.sc/); Experiment 2 was done via videoconference and the platform Real Eye (https://www.realeye.io). The sociodemographic questionnaire form was given to the parents in Google Forms via email. The ROWPVT-4: SBE test was applied during the videoconference using the test templates, while the experimenter's voice was being heard. Finally, the Flanker Task was also applied using the platform "Gorilla" (https://gorilla.sc/).

Because of these online conditions, the parents were instructed to use a computer (not a cell phone or tablet) placed on a stable surface and equipped with a functioning integrated camera; to be in a room with good, natural lighting; to make sure the child was as comfortable as possible at an adequate distance from the screen (as instructed by the platform RealEye when the experiment is running); and to hold the child's head when instructed. They were also told that they must be in a room with as few distractions as possible, abstain themselves from giving an answer or helping the child in any way, and that the child must be well-rested and fed before the session.

In addition, to have the best eye-tracking results, the platform Real Eye recommends the following system requirements:

1. Using only a computer or a laptop.
2. Having a webcam with 1080p @ 30FPS or 720p @ 60 FPS, to have good image quality.
3. Using the browsers Google Chrome or Microsoft Edge and having them updated. It is specified that the browser Firefox should not be used.
4. Having an updated Windows 10 or macOS X.

5. Processor 4-core Intel, 3 GHz or faster, of the 5th generation or later.
6. If using a laptop, having it plugged into the power cord.
7. Having a memory of 8 GB of RAM with at least 4 GB of memory that is not used or that is available.
8. Having a graphics card DirectX 12 with updated drivers. Having an integrated Intel Graphics of the 5th generation or later is enough.
9. Not having any other webpage or programs (other than the ones required for the experiments) open.

If these requirements were not met, the participants were not able to take part in the experiments.

2.8. Instruments and Materials

The materials and instruments that were used in this research were the following:

- **Informed consent:** A document that explains, in a clear manner, the specifications of the current investigation, describes the different experimental tasks and the role of the participant in them. It states the anonymity of the participants in the data that were analysed, asks their permission if they were recorded in any way, and specifies their right to leave the research or withdraw their results if they consider it necessary. In addition, it also contains information of the institution and the researchers involved in the investigation. This document must be signed by the tutors or parents of all participants that are voluntarily willing to take part in the experiments.
    a. **Sociodemographic questionnaire form:** A standardized questionnaire from the UNAM Psycholinguistics Laboratory, that the parents of the child

are asked to fill out. It asks for information about the infant, such as name, date of birth, age in months and days, gender, if the child was premature and in which weeks was she or he born if that is the case, if there were any difficulties during childbirth, the weight of the child when she or he was born, if the child has any health problems, if the child has any visual or auditory problems, if someone speaks to the child in another language (in which case, who and how often), the number of siblings and birth order of the child, the people that live with the child and the main caregiver of the infant at home, and finally, if the child goes to a day-care centre (and if so, how many hours a day and how many days a week). It also asks information about the parents, such as name, age, occupation, the number of years each parent has spent in every educational level and the state they live in. This questionnaire was administered using Google Forms using the following link: https://forms.gle/MJLbGbWRUXP68gEn9

- **Spanish-Bilingual Edition of the Receptive One-Word Picture Vocabulary Test-5 (ROWPVT-4: SBE):** According to Martin (2013), the ROWPVT-4: SBE is a receptive vocabulary test and verbal aptitude screening that covers a wide range of ages –from 2 to 70+ years of age–. It has different difficulty levels depending on the age of the participant, and raw scores can be converted into standard scores, percentile ranks, and age equivalents. In the present dissertation, in order to assess a language comprehension (specifically, child's vocabulary development) that is adequate for the child's age, age equivalents were used, which are compared to chronological ages. The test consists of 180 test plates with 4 colourful pictures in

each of them. The participant must select the picture that best describes a word that is said by the experimenter. Because of this, it has a simple and straightforward application, and pre-schoolers can easily answer by pointing, uttering their answer, or even nodding or shaking their heads while the experimenter points at each picture. The time of application consists, approximately, of 20 minutes, and the scoring can be completed in 5 minutes. It was validated in Spanish with Latin-American speakers, including Mexicans, in their countries and in the USA. It was also validated for its online application. The test was adapted online using the test plates provided in a Zoom (https://zoom.us/es) videoconference, where the four images were shown on a screen. The child selected the correct image saying its number on the test templates or pointing at the screen (in which case, the parent said the number of the image, knowing beforehand that they must not correct the child). Regarding this investigation, it was expected that the participants have a cognitive and language level that is typical for their age. Participants who did not pass this test were not considered for the experiment, as they had cognitive or language deficits that may affect the results. This test was only used to ensure that the participants did not have any language comprehension difficulties.

- **RealEye:** An eye-tracking online platform that was used to assess the gaze patterns and timing of eye-fixations of the participants on a screen as they listen to a series of utterances. This platform uses the computer's camera to track the participants' eyes. For this reason, good lighting is fundamental. According the the website (https://www.realeye.io/), RealEye is approximately 100 px (~1.5 cm) accurate, with an average error visual angle of ~4.17 deg. The prediction of the gaze point

has a frequency up to 60 Hz. The equipment and setup required for an optimal analysis is a Google Chrome or Edge (10) web browser, a laptop with an integrated camera or a USB webcam (wither Windows 10 or MacOs), and a screen resolution of 1024x968 pixels or more. The data given by the eye-tracker is exported in comma-separated values (.csv) files, which include information such as the participant's number, age and gender (if provided); test id, date of creation and duration in milliseconds (msec); the participant's browser width and height; and test raw data that includes the coordinates, in pixels, of the participant's eyes on the screen, and the time in msec in which that look took place.

- **Gorilla:** An online platform (https://gorilla.sc/) that provides tools for an online experimental design of behavioural research. Stimuli can be presented by images, videos, text, or audio, controlling for their duration and type of response (by writing, clicking on the image, pressing certain keys, or speaking). Instructions can be added freely. The data given is exported in comma-separated values (.csv) files, which include information depending on the type of response codded in the experiments. Voice recordings can also be downloaded if required.

2.9. Experimental Procedure of Experiment 1

During the videoconference, the experimenter started a warm-up phase, in which they introduced themselves to the child and showed them the experimental images, asking questions about what was above or below each image. This session helped the child to become familiar with the stimuli and sentence structure (especially since experimental stimuli have sentences with definite articles, and hence the images need to be introduced first). For this purpose, the combination of images was different from what they saw in the experiment itself, which

presented animate and inanimate objects separately. In this case, both animate and inanimate objects were presented at the same time on screen, as seen in *Appendix A*. This phase also helped to see if the child identified the concepts of "above" and "below" correctly which, according to the Norms for the MacArthur–Bates Communicative Development Inventories for Mexican Spanish-speakers (Jackson-Maldonado et al., 2003; Weisleder & Fernald, 2013), are concepts already known by 73% of Mexican Spanish speakers by 30 months of age.

With the same images, a practice session began, where the experimenter said correct or incorrect statements about the images, such as: *"¿La araña está arriba y la mano está abajo?"* ["Is the spider above and the hand below?"]. The experimenter then asked the child if this is correct or not. If the statement was incorrect, the experimenter encouraged the child not only to say if the statement was correct or not, but also to give the correct answer. For this purpose, intonation was not important, as this phase's goal was to introduce the child to the dynamics of the experiment. Because of this, recordings were not used, and the experimenter continued asking the child until they made sure the child knew how the experiment worked.

The instructions in the practice session were as follows: "*Ahora vamos a jugar un juego. Me voy a tapar los ojos y tú me vas a decir si me acuerdo bien de los dibujos. ¡Vamos a practicar! A ver si me acuerdo bien…*" ["We will now play a game. I will cover my eyes and you will tell me if I remember the drawings correctly or not. Let's practice first! Let's see, if I remember correctly…"].

After this, the experimenter gave feedback about the child's performance until the child understood the experiment three consecutive times (that is, after the child corrected the experimenter). Then, the experimental trials began. The experimenter told the child that they would now continue the game and instructed the parent to open the Gorilla link. The instructions

in this part were as follows: "*¡Lo hiciste muy bien! Ahora vamos a empezar el juego con los dibujos que viste. Tu papá/mamá/cuidador va a abrir el juego. ¿Listo? ¡Empecemos!*" ["You did great! The game will now begin with the drawings you saw. Your mom/dad/caregiver will now open the game. Ready? Let's begin!"]

The experimenter remained in the videoconference with their camera and microphone off in case there was any problem. This also had the purpose of encouraging the child to give the correct answer when an incorrect statement was said about the images, as complete answers were needed for the acoustic analysis of the pre-schoolers's responses. So, for example, if the recording stated: "The ant is above the owl, right?", and the answer was "no", the experimenter then opened their microphone and asked "Then, what is the right answer?" to ensure the child made a correction.

The parents were instructed not to help the child in any way and not to repeat the sentences; if needed, they had the option to press a "Repeat" button to listen to the utterance again. They also had to press a "Continue" button to advance to the next stimuli. This ensured that everything the child had to say was recorded.

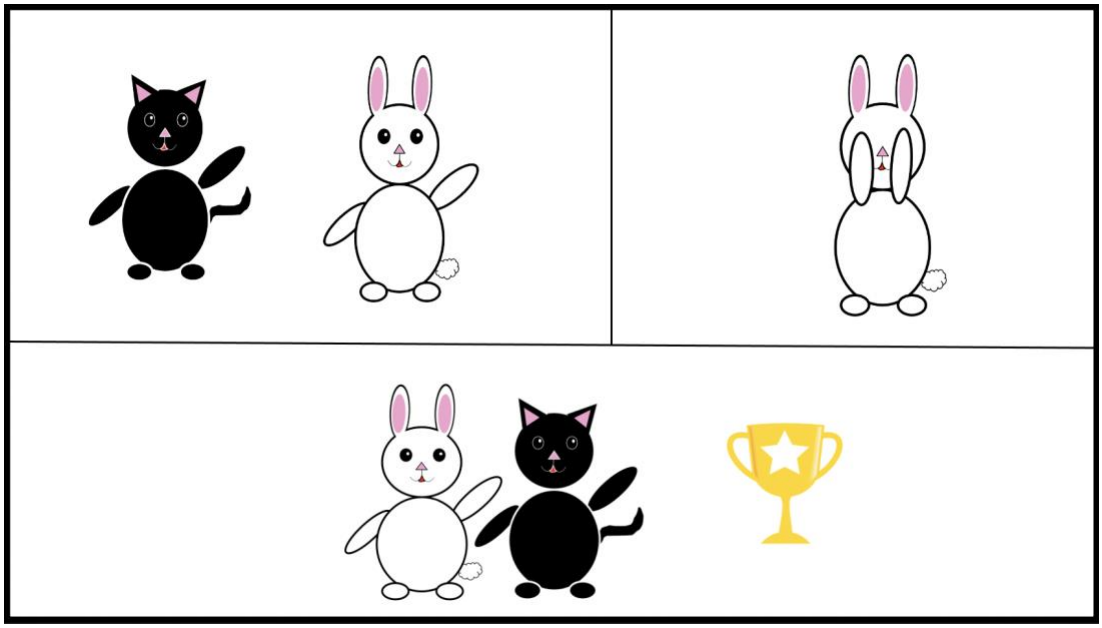
The experimental trials consisted of the 36 stimuli presented randomly on screen divided in 4 blocks with a pause between them where the child was congratulated for their effort. Before the presentation of the stimuli, an animated character of a cat appeared on screen, telling the child that she needed help for a friend, who needed to remember the images that the child saw earlier. Then, another character, this time a rabbit covering her eyes, appeared. This way, a conversational context was created in which focus was required to correct the rabbit. The instructions were as follows:

"*¡Hola, gracias por venir a jugar! Necesitamos tu ayuda para recordar los dibujos que viste. Mi amiga quiere saber si se los aprendió bien, ¿le ayudas? Se va a tapar los ojos y tú le dirás si recuerda bien los dibujos. Cada vez que le ayudes, aparecerá un premio. ¡Consigue todos! ¿Listo? ¡Vamos a empezar!*" ["Hi, thanks for coming to play! We need your help to remember the drawings you saw before. My friend needs help to know if she learnt them correctly, would you help her? She will cover her eyes and you will tell her if she remembers the drawings correctly. Every time you help her, a reward will appear. Get them all! Are you ready? Let's begin!"]

Instead of a fixation cross, the rabbit covering her eyes appeared on screen to remind the child that she was asking the questions. In addition, between each block, the characters appeared again on screen along with a cheerful melody. Then, a trophy was presented with the sound of applauses. After this, the next block began. This pattern was repeated until the experiment ended. At the end, the characters appeared again congratulating the child, saying: "*¡Lo hiciste muy bien, gracias por tu ayuda! ¡Hasta luego!*" ["You did great, thank you for your help! See you later!"]. For an example of the characters used, see Figure 6.

Figure 6

Example of characters used in Experiments 1 and 2



Note. Up, from left to right: presentation of the characters and fixation cross with the rabbit covering her eyes. Down: The characters presenting the first trophy out of four between the experimental blocks.

Results of the Pilot Test of Experiment 1 can be seen in *Appendix E*. The variables of the experiment are presented below.

2.9.1. Variables of Experiment 1

Condition (Independent Variable). This variable had two levels: SC and PPC; the participants were assigned to either of them. Six images of different animals or different objects were shown on screen to the participant. The experimental auditory stimuli had prosodic prominence in the subject (SC) or prepositional predicative (PPC) of the utterances.

Number of Adequate Corrections (Dependent Variable). The number of adequate corrections to the SC when the experimenter marked emphasis on the subject of the sentence, and the number of adequate corrections to the PPC when the experimenter marked emphasis on the prepositional predicative of the sentence.

Pitch tones (Dependent Variable). When given complete answers with the form SVO, intonational curves on the accented syllable of the subject and the prepositional predicative of the utterances were assessed using Praat (Boersma & Weenink, 2006), as well as the Sp_ToBI (Tones and Breaks Indices) transcription system (Beckman et al., 2002) to analyse the type of pitch used inside the phrase. Pitch height was also measured in semitones.

Boundary tones (Dependent Variable). When given complete answers with the form SVO, boundary tones inside the utterance were assessed using Praat (Boersma & Weenink, 2006), as well as the Sp_ToBI (Tones and Breaks Indices) transcription system (Beckman et al., 2002) and the prosodic hierarchy indices by Nespor and Vogel (1994) to analyse focus marking trough phrasing.

Duration (Dependent Variable). When possible, the duration of the accented syllable of the subject and the prepositional predicative of the utterances was assessed using Praat (Boersma & Weenink, 2006) as it may be a prosodic cue that marks focus on a word (Hualde, 2013). This duration was measured in miliseconds.

2.10. Flanker Task Procedure

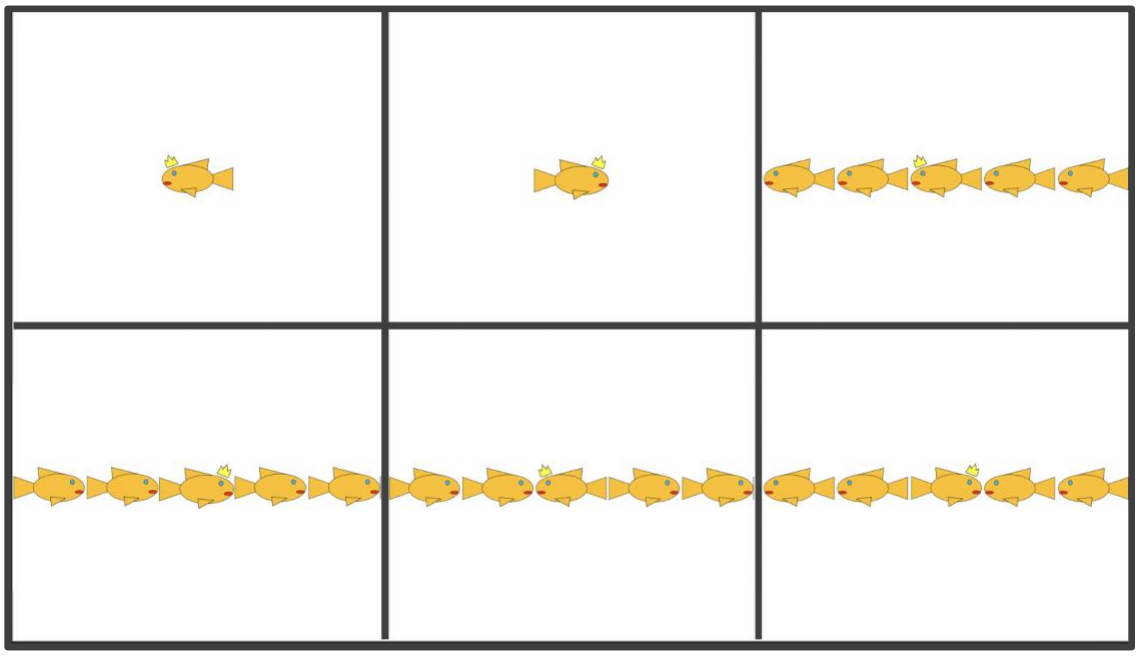
The Flanker Task that was presented is based on Rueda's et al. (2004) and Höhle's et al. (2016) designs, which are appropriate for 4-year-old pre-schoolers. Instead of arrows, as it is normally presented to adults, the task showed five orange fish aligned horizontally at the centre of the screen, facing left or right. The objective is to press a button, on the side the fish in the

middle (the target fish) is facing, as quickly as possible. In this case, the buttons were the keys “A” and “L” on the keyboard. Each had a paper with the colour “red” or “green” respectively. Such papers were taped to the keys by the parents prior to the experimental session as instructed by the experimenter.

There were three conditions: A Neutral Condition (NC) in which the target fish was alone, a Congruent Condition (CC) in which the target fish faced the same direction as the rest of the fish, and an Incongruent Condition (IC) in which the target fish faced the opposite direction than the rest of the fish. Each condition had two types of stimuli where the target fish faces left or right, resulting in six different stimuli. Each of these six stimuli were presented in eight experimental blocks, forming, in total, 16 trials per condition. In addition, two practice blocks, consisting of the six stimuli, were also presented at the beginning. Examples of each stimulus are presented in Figure 7 below:

Figure 7

Stimuli of the Flanker Task



Note. In order, from top to bottom and from left to right: NC left, NC right, CC left, CC right, IC left and IC right. Each square represents what is shown in the screen at each time.

The task was presented as a game about a fish looking for food (which only appeared at the end of each block). To avoid any confusion by children, the target fish wore a crown and was presented as "*El pez rey*" [The King Fish] when more fish were around. The task began with a recording of a female, Mexican Spanish-speaker telling the story about the King Fish being sad and needing the child's help to find its food.

The instructions were also written on screen in case the parent or caregiver needed to read them and began as follows: "*¡Vamos a jugar un juego! ¿Me ayudas a cuidar a un pececito? Él es el pez rey. Está triste. Necesita de tu ayuda para conseguir comida. El pececito aparecerá en medio de la pantalla. Cuando voltee hacia este lado (izquierda), presiona la tecla roja ('A')*

para decirle que hacia allá está su comida. Busca la tecla roja en tu teclado". [Let's play a game! Would you help me to take care of a little fish? This is the King Fish. It is sad. It needs your help to find its food. The little fish will appear on the center of the screen. When it turns to this side (left), press the red key ('A') to tell it that its food is in that direction. Search for the red key on your keyboard."]

Then, 10 seconds were given for the child to find the red key, which was on the left side of the keyboard, and to press it when the fish is facing left after the instruction: "*Presiona la tecla roja*" ["Press the red key"]. An animation with music was then played out in which the fish moved to that side of the screen. When it arrived, the child was congratulated. The same happened with the green key ("L"), that must be pressed when the fish is facing right, with the instruction: "*Cuando voltee hacia este otro lado (derecha), presiona la tecla verde para decirle que hacia allá está su comida. Busca la tecla verde en tu teclado. Presiona la tecla verde*". [When it turns to this side (right), press the green key to tell it that its food is in that direction. Search for the green key on your keyboard. Press the green key."]

Before any block, a fixation cross appeared on the screen for 1 second. After this, a first practice block with the six stimuli in the order shown in Figure 7 was presented. The instructions given before the block began were the following: "*¡Muy bien! Si vienen otros peces, sólo alimenta al pez rey. ¡Practiquemos!*" ["Well done! If other fish come, only feed the King Fish. Let's practice!"]

Children received feedback for a correct or incorrect answer: A bell sound was played when the answer was correct, and a buzzer sound was played when the answer was incorrect. Time was given for the child to familiarize with the stimuli. When this block ended, a second practice block began in which the child was instructed to answer as fast as possible this time. If the child

did not answer after 4 seconds, the buzzer sounded, and the answer was marked as incorrect. The instruction was: "*¡Muy bien! Ahora contesta lo más rápido posible. Recuerda: sólo hazle caso al pez rey. Si él solito voltea para acá (izquierda) presiona rojo. Y si él solito voltea para acá (derecha) presiona verde. No le des su comida a los demás. ¡Empecemos!*". ["Well done! Now answer as quickly as possible. Remember: you must only pay attention to the King Fish. If it is the only one facing this direction (left) press red. And if he is the only one facing this direction (right) press green. Do not give food to the rest of the fish. Let's start!"]

When the practice blocks ended, the eight experimental blocks began. To keep the children's attention, an instruction video was first displayed in which the child was invited to help the fish eat eight apples, which were presented on screen. The instructions were as follows: "*¡Muy bien! Es momento de ayudar al pecesito. Cada vez que presiones el botón correcto, estará más cerca de su comida. Cada vez que lo acerques, aparecerá más comida. ¡Haz que se coma todas! ¿Listo? ¡Empecemos!*" ["Well done! It is time to help the little fish. Every time you press the correct button, it will be closer to its food. Every time it is closer thanks to you, more food will appear. Help it eat all of them! Ready? Let's begin!"]

Then, the experimental trials began. At the end of each block of six stimuli in randomized order, an animation of the fish eating one of the apples was displayed. At the end of the experiment, the last apple is eaten, and the child was congratulated, telling them the fish is now happy and not hungry.

The Pilot Test of the Flanker Task can be consulted in *Appendix F*; furthermore, individual results from the Flanker Task of the experiments can be seen in *Appendix G*.

2.10.1. Variables of the Flanker Task

Condition (Independent Variable). This variable had two levels: The CC, where the target fish faces the same direction as the rest of the fish, and the IC, where the target fish faces the opposite direction to the rest of the fish.

Interference Rate using Reaction Times (RT) (Dependent Variable). Interference rates was measured by subtracting the RT of incongruent trials to the RT of congruent trials. This is a measure of inhibitory control ability, as a less developed inhibition is shown in the latency of the response and a major error rate (Höhle, 2016).

2.11. Results of Experiment 1

2.11.1. Participants of Experiment 1

For data analysis of Experiment 1, one of the participant's recordings was inaudible, and hence was excluded from this analysis. This left a total of 40 participants, randomly assigned to each condition (SC, *n*= 20; PPC, *n*= 20).

2.11.2. Analysis Criteria of Experiment 1

The first analysis consisted in determining the percentage of correct responses to the stimuli presented. A response was classified as "correct" only if the participants answered with a correction to the experimenter (for example, "No, the squirrel is the one that's above the owl"), a plain "no", or another negative answer to the experimental stimuli; and "yes", "correct" or another affirmative answer to the control stimuli.

In the second analysis, responses to the experimental stimuli were analysed, where it was expected that the pre-schoolers corrected the subject in the SC (for example, considering that the owl is above the squirrel in an image, correct that the owl is above the squirrel in "The SPIDER is above the squirrel, right?") and the prepositional predicative in the PPC (in this case,

considering that the frog is below the spider in an image, correct that the frog is below the spider in “The spider is above the SQUIRREL, right?”). In this case, a Chi-squared test was performed to test if there is a relation between the corrections to the subject and prepositional predicative and both conditions.

An acoustic analysis was then performed with sentences that had the same syntactic structure in the SC and PPC. Its purpose was to examine if there was an acoustic difference between them, as it was perceived that, even when the answers correcting the prepositional predicative, which were the majority, were the same in both conditions, emphasis was made by the participants on the subject in the SC and on the prepositional predicative in the PPC. However, responses to the subject were also analysed to make a comparison. These recordings had answers where the pre-schoolers corrected the experimenter uttering a sentence with subject, optionally a verb, and prepositional predicative configuration (for example, “No, the squirrel is above the owl”) in order to compare emphasis on the subject and prepositional predicative across conditions.

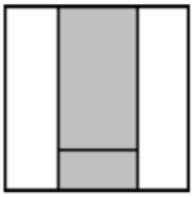
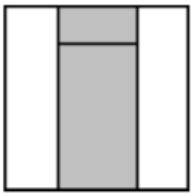
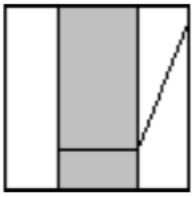
For this acoustic analysis, which was also done with the experimental stimuli, the transcription model that was used was the Sp_ToBI (“Spanish Tones and Breaks Indices”, based on the original version,“Tones and Breaks Indices”, by Beckman et al., 2002) in its latest version (Hualde & Prieto, 2015), which has used before for the analysis of focus production in preschool Mexican Spanish speakers (Villalobos-Pedroza, 2021). Likewise, the notations established by Jun (2005) were used, where the symbol “>” indicates prealignment of the peak of the intonational curve inside the tonic syllable, while “<” represents the peak of the tonal accent on the posttonic syllable, both notations in the case of bitonal accents. Furthermore, following Martín Butragueño (2019), the upstep diacritic (“¡”) is used to represent an ascent greater than three semitones (st); additionally, the downstep diacritic (“!”) is used when the tone is lower by 3

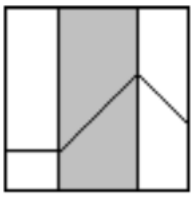
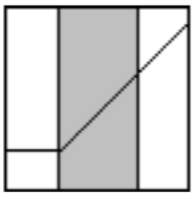
st. from the rest of the utterance. In this initial approach, a qualitative criterion was followed for pre-alignment marking.

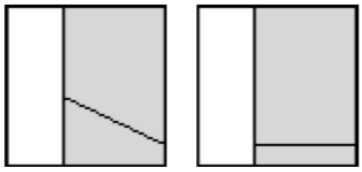
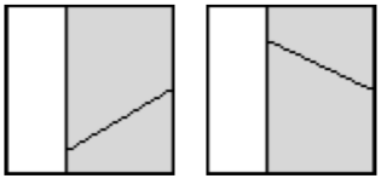
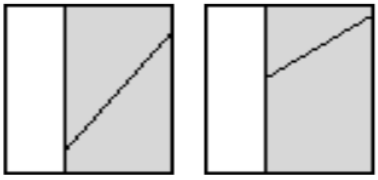
This model is based on the Autosegmental Metrical model (AM), which analyses the tones (pitch accents) and breaks (pauses or interruptions of the phrase due to a higher or lower pitch in the boundary of a phrase) in spoken language. In the case of Spanish, these tones are analysed on the accented syllables. The pitch on tones and phrase boundaries is classified as high (H) or low (L), adding the symbol "*" on the accented syllable, and the symbols "-" on a boundary inside the utterance and "%"at the end of the utterance. In addition, when an elevation or decline of the melodic curve is perceived inside the accented syllable, a bitonal pitch is presented adding the symbol "+" between two tones, forming combinations such as "L+H*" or "H+L*", for example. A complete characterization of the pitch accents and pitch boundaries is presented in Table 5, adapted from De-la-Mota, Martín-Butragueño and Prieto (2010) and Villalobos-Pedroza (2021):

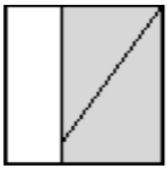
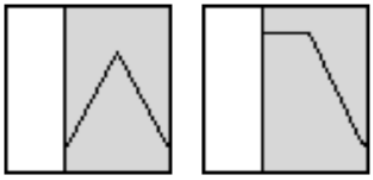
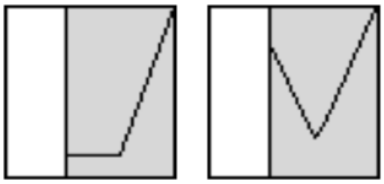
Table 5

Pitch Accents and Pitch Boundaries Used for the Transcriptions

	Scheme	Transcription	Description
Monotonal pitch accents		L*	Progressive fall or maintenance of a low value of the F0 on the tonic syllable.
		H*	High and maintained F0 on the tonic syllable.
		¡H*	A high pitch that is not the highest in the utterance but is still high.
Bitonal pitch accents		L*+H	Low pitch on the tonic syllable with a decline of the F0 on the posttonic syllable.
		L+H*	Rise of the F0 that is shown at the beginning of the

Scheme	Transcription	Description
		tonic syllable with its peak being reach at the end of it.
	L+¡H*	If the rise of the L+H* is higher than 3 semitones.
	L+<H*	Progressive rise of the F0 during the tonic syllable with its peak on the posttonic syllable.
	L+>H*	Rise of the F0 where its peak is reached within the tonic syllable, in its middle or before it ends. The rise is higher than 3 semitones.
	H+L*	Fall of the F0 inside the tonic syllable.

	Scheme	Transcription	Description
Monotonal boundary tones		L% (end of the utterance) L- (end of a phrase inside the utterance)	Fall of the F0 from a high point on the tonic syllable (the fall must be of 1.5 semitones or higher), or a maintained low pitch.
		!H% (end of the utterance) !H- (end of a phrase inside the utterance)	Fall or rise of the F0 that is lower than 1.5 semitones, or a maintained pitch that comes from a high F0 on the pretonic syllable.
		H% (end of the utterance) H- (end of a phrase inside the utterance)	Rise of the F0 on a tonic syllable that comes from a lower pitch.

	Scheme	Transcription	Description
		¡H% (end of the utterance) ¡H- (end of a phrase inside the utterance)	Rise of the F0 that comes from a low or high pitch which rise is higher than a tone H% or H-
Bitonal boundary tones		HL% HL-	Rise and fall of the F0 after a low nuclear accent or the maintenance of the F0 followed by a fall.
		L!H% L!H-	Fall and rise of the F0 after a high nuclear accent or the maintenance of the F0 in a low pitch, followed by a rise.

Note. In the scheme, the divisions of the box represent the syllables. The stressed syllable is grey, and the lines represent the melodic curve and its movements in such syllables. The *tonal peak* refers to the maximum point of the syllable.

In addition, phrasing was transcribed using the Sp_ToBI transcription indices, which in turn are equivalent to the levels of the prosodic hierarchy of Nespor and Vogel (1994):

- 0= When two words form a single prosodic word because one of them does not have an accented syllable. For example, “the” and “spider”, where “the” is an unaccented word and hence is perceived as a single prosodic word with “spider”: [ðə’spaɪdə͡ɹ].
- 1= A limit between two prosodic words (for example, [ðə’spaɪdə͡ɹ] and [rʌns] in the sentence “The spider runs”).
- 2= When a higher pitch inside the utterance is perceived as a pause inside of it (for example, in the sentence “The SPIDER runs”, if the last syllable of “spider” has a high pitch, this prominence may be perceived as a separation between this word and the rest of the phrase).
- 3= A pause inside the utterance (for example, the comma in “Then, the spider ran”).
- 4= The end of the utterance.

The analyses using these transcriptions were done using Praat (Boersma & Weenink, 2006) with a pitch range of 75 to 700 Hz to be able to visualize the high pitches produced by children of those ages. Additionally, the pitch was measured in st because this measure is used in previous studies with acoustic analysis, is not sensitive to gender differences and is more representative of acoustic perception, as stated by Pamies-Beltrán et al. (2001). When two tones are perceived inside the accented syllable, the criteria of 1.5 st was used to mark bitonal pitches, as it is the minimal number of semitones required to perceive a change in pitch, as reported before (Pamies-Beltrán et al., 2001).

In addition, pitch height, elevation in semitones of the melodic curve (or pitch elevation) and the duration of the accented syllable of the subject and prepositional predicative of the sentence in both conditions was analysed to determine differences in the emphasis used by the participants. The statistical analysis used was a Mann-Whitney *U* test in order to compare differences between conditions, as data was not normally distributed.

2.11.3. *Results from verbal responses and acoustic analysis*

A total of 1,440 audio files (720 per condition) were obtained. From this total, only the recordings that belonged to the control stimuli (where the participants were expected to answer with "yes" or "correct" to the correct statements about the images) and to the experimental stimuli (where the participants had to correct the wrong statement said by the experimenter) were considered for the analysis. This left a total of 960 audio recordings (480 per condition).

There was a total of 96% of correct responses to the SC and a total of 95% correct answers to the PPC. In consequence, the percentage of correct answers was above chance in both cases, which indicates that the participants understood the task correctly.

Comparing the answers given to the experimental stimuli in both conditions, where pre-schoolers had to correct the heard statements, the responses given by the participants showed that, in the SC, there was a total of 49 corrections to the subject and 160 corrections to the prepositional predicative (31 more were errors or blank recordings). In contrast, the PPC had 74 corrections to the subject and 141 corrections to the prepositional predicative (25 more were errors or blank recordings).

Although there was not a significant 2x2 association between condition and corrected word ($x^2(1) = 2.969$, $p= 0.085$) with a small effect (Cramer's $V=0.085$), these results are the opposite pattern to what was expected: there was a greater number of corrections to the subject in the PPC

than in the SC. However, there were more corrections, in general, to the prepositional predicatives of the sentence, resulting in identical answers in both conditions.

2.11.3.1. Acoustic analysis to the corrections to the prepositional predicative. An acoustic analysis was performed in these sentences where the prepositional predicative was corrected, which had the same syntactic structure in the SC and PPC. For this, 152 sentences were selected in the SC and 126 in the PPC (from the total of 240 per condition). A comparison on the tone pattern was made first between the subject and the prepositional predicative of the utterances in both conditions. The results showed a different tone pattern between conditions regarding the subject of the sentence, as shown in Table 6 below:

Table 6

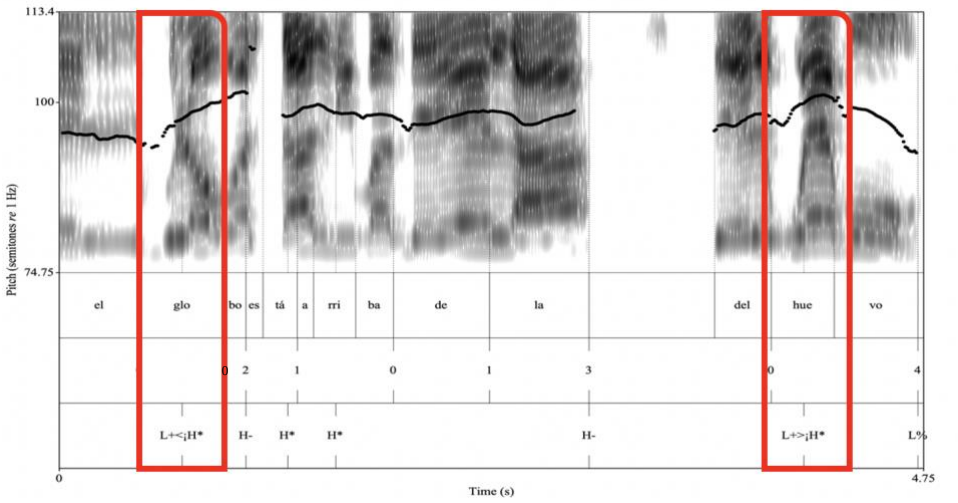
Pitch Accents Observed in the Responses Given by Pre-schoolers in Experiment 1

Pitch accents on the subject		**Pitch accents on the prepositional predicative**	
Subject Condition	Prepositional Predicative Condition	Subject Condition	Prepositional Predicative Condition
L+<H* (29.80%)	L+<H* (22.22%)	L+>H* (28.47%)	L+>¡H* (23.01%)
H* (26.49%)	H* (20.63%)	L+>¡H* (19.20%)	L+>H* (19.04%)
L+<¡H* (16.55%)	L+<¡H* (19.04%)	H* (15.23%)	H* (15.87%)
L+>H* (7.28%)	L+>¡H* (13.49%)	L* (11.25%)	L+<H* (11.11%)
L* (5.29%)	L+>H* (11.11%)	L+<H* (7.94%)	L* (7.93%)
L+H* (3.97%)	L+¡H* (4.76%)	L+H* (5.96%)	L+<¡H* (7.14%)
L+¡H* (3.31%)	L+H* (3.96%)	L+¡H* (5.29%)	¡H+L* (5.55%)
L+>¡H* (3.31%)	¡H+L* (1.58%)	H+L* (2.64%)	L+H* (3.96%)
H+L* (2.64%)	H+L* (1.58%)	L+<¡H* (2.64%)	H+L* (3.17%)
¡H+L* (0.66%)	L* (1.58%)	¡H+L* (1.32%)	L+¡H* (3.17%)

Table 6 shows that in both conditions, the most used tone pattern on the subject of the utterance was L+<H*, which is usually used in Spanish to mark the topic of the sentence (Hualde, 2013); which means it is not focused. This was not the case of prepositional predicative in both conditions, where the most common pattern on the subject was L+>H* (L+>¡H* in the PPC, which only means that f0's raise was greater), usually used to mark contrastive focus in Spanish (Hualde, 2013). The most common tone patterns of both conditions, shown in Table 6, can be seen in Figure 8 and Figure 9.

Figure 8

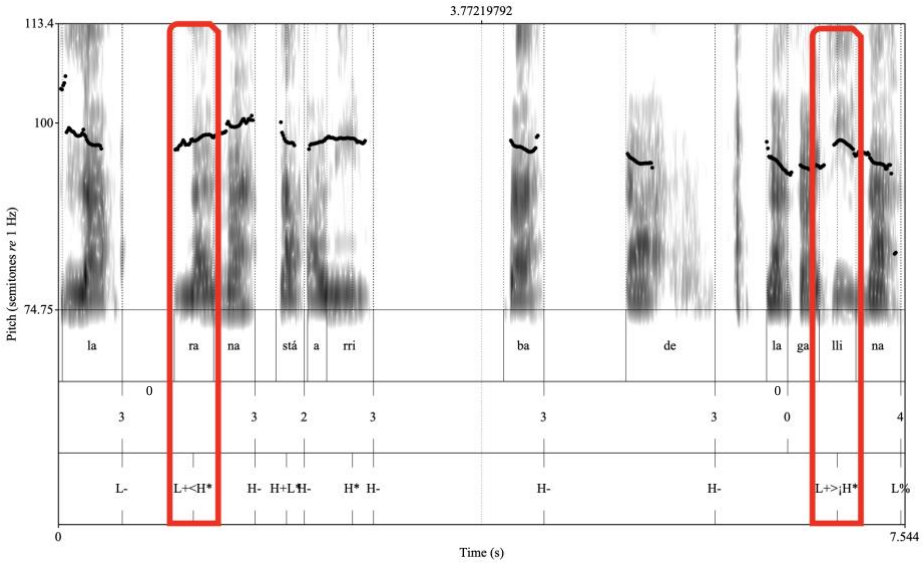
Example of Prepositional Predicative Corrections Made by Pre-schoolers in the Subject Condition



Note. Children heard the utterance: "The BALLOON is above the hand, right?" and replied: "No, the balloon is above the egg." Tonic syllables of the subject and prepositional predicative are marked in red.

Figure 9

Example of Prepositional Predicative Corrections Made by Pre-schoolers in the Prepositional Predicative Condition

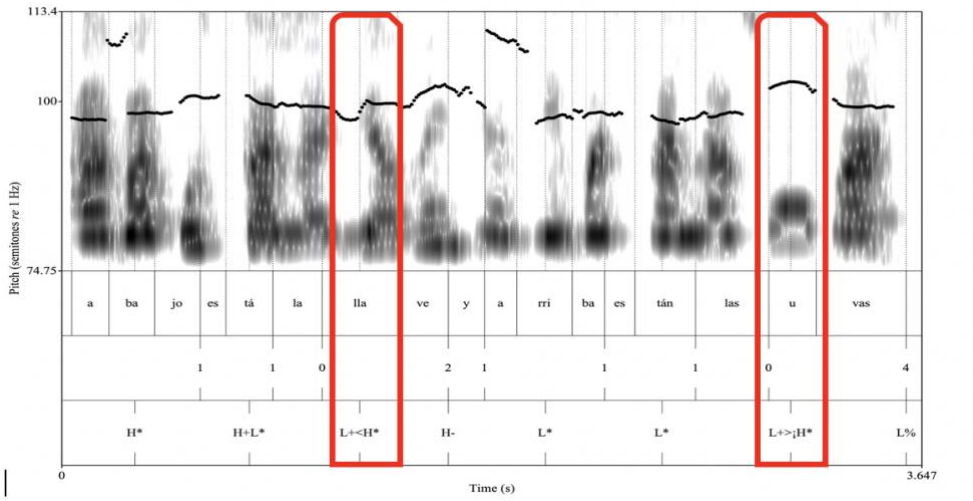


Note. Children heard the utterance: "The ant is above the HEN, right?" and replied: "No, the frog is above the hen." Tonic syllables of the subject and prepositional predicative are marked in red.

Regarding sentence structure, in the PPC it was more common for participants to make syntactic movement to place the subject at the end of the utterance (as seen in Figure 10, where "grapes" is the subject); here, a tone pattern of L+>¡H* was more common on the subject and a tone pattern of L+<H* was more common in the prepositional predicative. This happened in 22% of the cases, while in the SC it happened only in 7% of the cases.

Figure 10

Example of Prepositional Predicative Corrections Made by Pre-schoolers in the Prepositional Predicative Condition with Syntactic Movement



Note. Children heard the utterance: "The grape is above the EGG, right?" and replied: "No, the key is below, and the grapes are above." Tonic syllables of the prepositional predicative and subject are marked in red.

Furthermore, the pause between the first element of the utterance (subject or prepositional predicative, depending on syntactic movement) and the rest of the sentence in is present both conditions in most of the responses, as an actual pause (Figure 9) or a sudden drop of the intonational curve that is perceived as a pause (Figure 8 and Figure 10).

In addition, the pitch height, the elevation in semitones of the melodic curve (or pitch elevation) and the duration of the accented syllable of the subject and prepositional predicative of the sentence in both conditions were analysed. The results showed a significant difference between conditions regarding the subject's pitch height (U= 7778, p<0.011) and pitch elevation (U= 7248, p<0.001). The subjects of the PPC had a greater pitch height ($\bar{x}$= 101.64 st, SD= 9.36)

and pitch elevation ($\bar{x}$= 3.55 st, *SD*= 8.54) than those in the SC (height: $\bar{x}$= 99.92 st, *SD*= 2.87; pitch elevation: $\bar{x}$= 2.28 st, *SD*= 1.48). No significant difference was found between conditions in the duration of the subject (*U*= 8241, *p*= 0.067); nor in the prepositional predicative's pitch height (*U*= 8546, *p*= 0.145), pitch elevation (*U*= 8724, *p*= 0.235) or duration (*U*= 9388, *p*= 0.851).

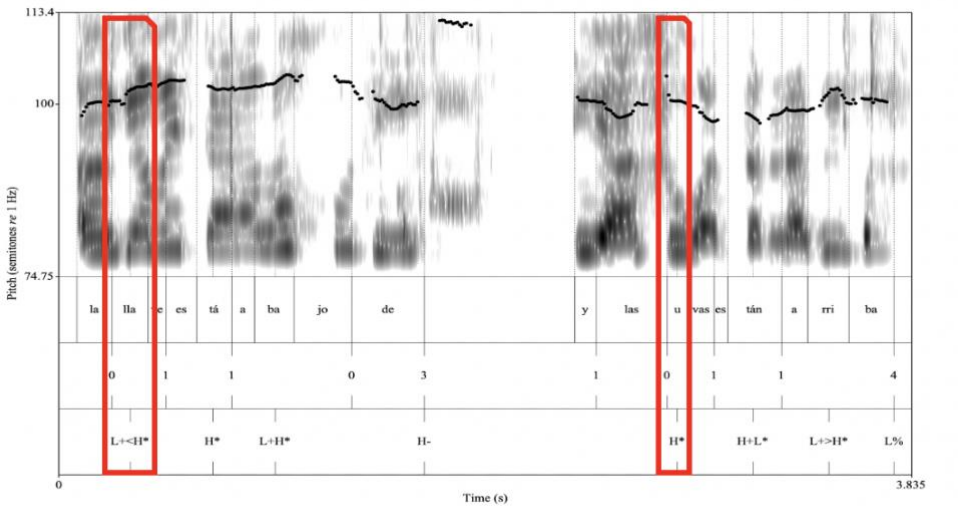
2.11.3.2. Acoustic analysis to the corrections to the subject. An additional acoustic analysis was done to those responses were pre-schoolers corrected the subject of the sentence in both SC and PPC, following the same criteria (45 of the SC and 66 of the PPC, from a total of 240 per condition). Even if they were the minority of responses, the fact that they exist may be related to pre-schoolers who had more developed cognitive abilities (Szendröi, 2004), which will be analysed in the following section with the results of the Flanker Task.

Analysis showed that the most common tone patterns on the subject in both conditions were not related to focus (Hualde, 2013). In both the SC and PPC, the most common tone pattern was H* (SC: 22.22% of the cases; PP: 18% of the cases). However, in PPC, the most common tone pattern was also L+<H* (20% of the cases), which may be interpreted as the topic of the utterance.

Nevertheless, it is important to point out that, in the more than half of the responses (51%), the subject was placed at the end of the sentence in the SC, which means it was interpreted as the focused word of the utterance (Gutiérrez-Bravo, 2008). An example of this is shown in Figure 11, along with the most common tonal patterns found. Furthermore, pauses were used as a mean of emphasis after the subject in 27% of the cases where the subject was not placed at the end, which is also a way of marking focus (Dorta-Luis, 2008). This also the case for the PPC, but with less than half of the cases (40%) in which the subject was placed at the end of the sentence.

Figure 11

Example of the Subject Corrections Made by Pre-schoolers in the Subject Condition

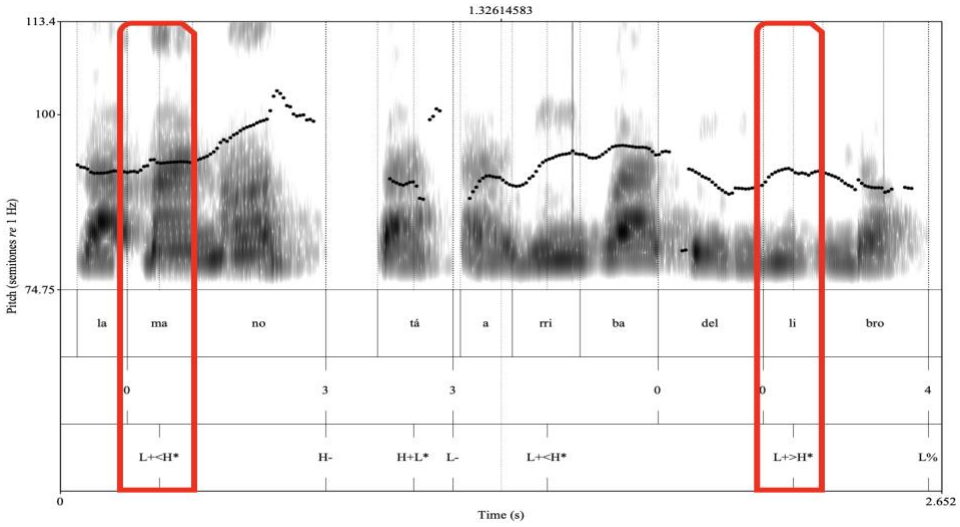


Note. Children heard the utterance: "The book is above the KEY, right?" and replied: "No, the key is below… and the grapes are above." Tonic syllables of the prepositional predicative and subject are marked in red.

Finally, regarding the prepositional predicative, the most common tone pattern in the SC was L+<¡H*, which accounts for a form of topic in Spanish, as explained before (Hualde, 2013). In the PPC, in contrast, the most common tone pattern was L+>H* (28% of the cases), which does accounts for a form focus (Hualde, 2013). This was the only difference found between conditions. An example of the results of the PPC can be seen in Figure 12.

Figure 12

Example of Subject Corrections Made by Pre-schoolers in the Prepositional Predicative Condition



Note. Children heard the utterance: "The key is above the BOOK, right?" and replied: "No, the hand is above the book" Tonic syllables of the prepositional predicative and subject are marked in red.

In comparison with the corrections to the prepositional predicative, the corrections to the subject show similar tone patterns for both the subject and prepositional predicative of the utterances. However, the subject is more emphasised with syntactic movement in the SC and the prepositional predicative is more emphasized with L+>H* in the PPC.

2.11.4. Correlation of the results of Experiment 1 with the results of the Flanker Task

For this analysis, only pre-schoolers that took part in both experiments (1 and 2) were considered. As stated before, one of the participant's recordings was inaudible, hence their data was excluded from Experiment 1. In addition, the data from three participants were excluded from Experiment 2 because they had a low percentage of data. The data of an additional

participant was excluded because of a sampling error, as they were wrongly assigned in both conditions. This left a total of 36 participants for this analysis (SC, n= 18; PPC, n= 17).

With respect to the analysis criteria, because data was not normalized, Spearman's correlation coefficient was used to find correlations between the proportion of verbal responses correcting the subject, the proportion of verbal responses correcting the prepositional predicative, and interference effect.

The results showed that only in the SC, a marginally significant correlation was found between the interference effect and the proportion of verbal responses to the subject (r (17)= 0.471, p= 0.049), while a marginally insignificant correlation was found between the interference effect and the proportion of verbal responses to the prepositional predicative (r (17)= -0.455, p= 0.058). This indicates that the greater the interference effect, the greater proportion of verbal corrections to the subject.

In the PPC, no significant correlations were found between the interference effect and the proportion of verbal responses to the prepositional predicative (r (16) = -0.056, p= 0.830) or between the interference effect and the verbal responses to the subject (r (16) = 0.056, p= 0.830).

2.12. Discussion of Experiment 1

Experiment 1 aimed to assess participants' comprehension of prosodic focus in utterances, specifically focusing on the subject (SC) or prepositional predicative (PPC), through verbal responses. Discrepancies between the images' positions in visual stimuli and the audio stimuli were introduced to evaluate participants' ability to correct the emphasized element.

In cases where the prosodic focus was on the subject (SC), the correct response involved correcting the subject, whereas for prepositional predicative focus (PPC), the correct response was to correct the prepositional predicative. Our hypothesis, based on studies in other languages,

suggested a greater number of correct verbal responses for PPC compared to SC (Szendröi, 2004; Höhle et al., 2016; Szendröi et al., 2018).

The results supported our hypothesis, revealing a higher number of responses to the prepositional predicative than to the subject. This tendency may be attributed to the common placement of the focused element at the end of the utterance, making its correction a default response. Notably, the use of syntactic movement in Spanish contributes to this pattern, with the focused element often positioned at the end (Hualde, 2013).

However, in these corrections to the prepositional predicative, a difference of emphasis to the subject or prepositional predicative could be perceived between conditions. Because of this, we performed an acoustic analysis. This analysis did not reveal significant differences in tone patterns between conditions. Both SC and PPC conditions exhibited specific tone patterns, with the subject marked as the sentence topic and the prepositional predicative marked for contrastive focus.

These lack of significant differences in tone patterns between conditions contradicted the results of the Pilot Test (see *Appendix E*), where there was a clear difference between conditions. The differences could be attributed to the presence of a pause in the auditory stimuli, a cue that was eliminated in the final experiment. Interestingly, the pause was found to be a reliable cue for focus, particularly among younger Spanish-speaking toddlers. This is the prosodic cue that they use more consistently than others, such as tone pattern or duration of the tonic syllable (Villalobos-Pedrosa, 2021). The population of this dissertation is older than the one in Villalobos-Pedrosa's (2021) study (4 to 5 years of age); however, it is possible that the pause between the subject and the rest of the utterance is still a clearer cue for focus for children of this ages. This possibility could be explored in future investigations.

But what happens with the lesser number of pre-schoolers who corrected the subject instead of the prepositional predicative? According to Szendröi (2004), even if most corrections are to the prepositional predicative, it is important to consider that some children do perceive contrastive focus differently.

An acoustic analysis demonstrated that the subject corrections were marked as contrastive focus. Syntactic movement in the PPC also indicated a potential interpretation of the subject as the focus, supported by a greater pitch elevation in the subject compared to SC. This indicates that some pre-schoolers that corrected the prepositional predicative in the PPC still could have interpreted the subject as the focus of the sentence.

As Szendröi's (2004) study suggests, these pre-schoolers may have more developed cognitive abilities (such as working memory). In this dissertation, working memory was not considered as a factor that may influence children's responses. However, considering the study of Höhle et al. (2016), it is possible that inhibitory control is also involved, especially when pre-schoolers must give a verbal response, in which case it might be harder for some of them to inhibit a standard answer (that is, a correction at the end of the utterance).

In this sense, a significant correlation was found between the interference effect and the proportion of verbal responses to the subject; however, this correlation was positive, which means that the greater the interference effect (that is, a less developed inhibitory control reflected in longer RT in in the Incongruent Condition), the greater the proportion of verbal corrections to the subject. This result must be taken with precaution, as it was marginally significant. Even so, the emphasised element (the subject, in this case) could be interpreted as the topic of the sentence. Because of this, participants usually make corrections to the prepositional predicative.

So, it is possible that pre-schoolers that corrected the subject, and hence had a less developed inhibitory control, could not inhibit an interpretation where the prepositional predicative is usually the one carrying emphasis (and, in this case, being the topic of the sentence). Consequently, that is the reason why they corrected the subject. This makes sense considering that the emphasised element is the one that tends to stay in their memories because of its salience, becoming the topic, and so, they correct the element without emphasis. This could be further explored in Experiment 2 using eye-tracking.

2.13. Experimental Procedure of Experiment 2

For this experiment, the instructions were different. In the practice trials, they were as follows: "*Ahora vamos a jugar un juego. Te voy a enseñar otra vez los dibujos, yo me los aprendí todos. Voy a taparme los ojos y me vas a decir si me los aprendí bien. Sólo puedes decir "sí" o "no", ¿está bien? ¡Practiquemos!*" ["We will now play a game. I will show you the drawings again, I already know all of them. I will cover my eyes and you will tell me if I remember them correctly. You can only say "yes" or "no", okay? Let's practice first!"]

After the practice trials ended, the instructions were the following: "*¡Lo hiciste muy bien! Recuerda estar muy calladito/a y prestar mucha atención. Ahora vamos a empezar el juego con otros dibujos diferentes. Tu papá/mamá/cuidador va a abrir el juego. ¿Listo? ¡Empecemos!*" ["You did great! Remember to be really quiet and pay a lot of attention. The game will now begin. Your mom/dad/caregiver will open the game. Ready? Let's begin!"].

Then, a *Gorilla* link was provided with the instructions of the task. In this case and with the purpose of keeping the child's attention on the screen without speaking, a narrative was created in which the infant was told that they have superpowers and the characters on screen (the same

as in Experiment 1) will be able to read their minds when they give an answer. The instructions given by the animated characters were as follows:

"*¡Hola, tenemos una misión importante para ti! A partir de ahora, ¡tienes superpoderes! Si te concentras y miras fijamente la pantalla, ¡podrás decirnos cosas con tu mente! ¡Necesitamos tu ayuda para recordar los dibujos que viste antes de empezar! Mi amiga quiere saber si se los aprendió bien, ¿le ayudas? No hables, quédate muy quieto y mira la pantalla. Con tus superpoderes nos dirás la respuesta correcta con tu mente cuando ella se tape los ojos. Quédate muy quieto, callado y mira la pantalla. Cada vez que observes en silencio, ganarás un premio. ¡Consíguelos todos! ¿Listo? ¡Vamos a empezar!*" ["Hi, we have an important mission for you! As of now, you have superpowers! If you focus on the screen, you can tell us things with your mind! We need your help to remember the pictures you saw before we started! My friend wants to know if she learned them well, can you help her? Don't talk, stay very still and look at the screen. With your superpowers you will tell us the correct answer with your mind when she covers her eyes. Stay very still, very quiet and look at the screen. Every time you observe in silence, you will win a prize. Get them all! Ready? Let's start!"]

After the instructions were presented, the platform *Gorilla* redirected the child to the platform *RealEye*. Then, the experiment began after a 4-pount calibration and adjusting the position of the infant's head on the screen. If the child moved, the experiment paused, and the parents had to align the participant's faces again to the outline of their eyes given by the platform, to ensure they were in same position as before. Additionally, between blocks and after congratulating the child for their progress, the characters reminded them to be quiet and still.

Because of the need of the child's head needed to be still, the parent's assistance was required for this purpose; they were instructed to hold the child's head softly throughout the experiment if necessary.

Results from the Pilot Test of Experiment 2 can be consulted in *Appendix H*. The variables of the experiment are presented below.

2.13.1. Variables of Experiment 2

Condition (Independent Variable). The same SC and PPC variables as before. This variable had two levels: SC and PPC; the participants were assigned to either of them.

Gaze patterns (Dependent Variable). The gaze direction can determine if the tonal prominence used in prosodic focus is correctly interpreted (Ito, 2018). For this reason, the looking patterns of the participants were assessed while they listened to recorded utterances. It was expected that when they listened to prosodic focus marked on the subject of the sentence, their eyes would go to the subjects on the screen first, and then to the prepositional predicatives. In contrast, if the prosodic focus was on the prepositional predicative of the sentence, it was expected that the participants would look at the prepositional predicatives first. It was also expected that they would compare the elements that were being contrasted, looking at the two options that showed contrast and not to the third option that was not mentioned in the utterance (which were distractors).

2.14. Results from Experiment 2

2.14.1. Participants of Experiment 2

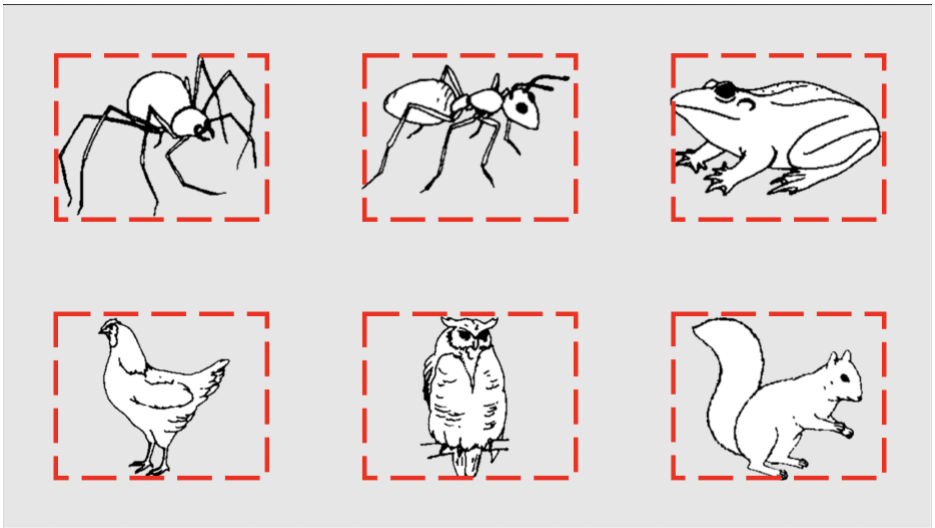
The same participants took part of this experiment. Two were eliminated from the SC (*n*= 19), and one from the PPC (*n*= 18) because they had a low percentage of data, which will be further explained in the following section. This left a total of 37 participants for this experiment.

2.14.2. Data processing and analysis criteria

The Areas of Interest (AoI) consisted of the six images presented (an example is seen in Figure 13); all of them of an equal size of 440x340 px in a 1920x1080 px background. Two images were classified as either the subject or the prepositional predicative of the sentence (for example, considering Figure 13, "The ant is above the squirrel", the ant is the *subject* and the squirrel, the *prepositional predicative*), two images were classified as competitor images, as they were contrasted in the utterances presented (following the example presented before, the frog would be the *subject competitor* and the owl the *prepositional predicative competitor*, as the ant is actually above the owl and the squirrel is actually below the frog), and the last two images were classified as distractors, as they were not considered in the utterance (in this case, the spider and the hen).

Figure 13

Areas of Interest of Experiments 1 and 2



Note: This image example was presented with the utterance "The ant is above the squirrel." AoI are represented with red lines, which were not part of the experimental image.

The window of analysis was the whole utterance in both conditions. However, differences between conditions are expected from the onset of the prepositional predicative of the sentence to the end of the utterance (in the SC, from 2880 to 5812 ms; and in the PPC, from 2500 to 5532 ms). In this window, comparisons between images to find the correct answer were expected to happen, as the participants needed to hear the prepositional predicative of the sentence to know if the statement was correct or not (for example, in the sentence “The spider is above the owl” they had to hear “owl” to know the correctness of the utterance). A timeline of the auditory stimuli presentation is shown in Figures 14 and 15.

Figure 14

Timestamps of Sentence Sections of Auditory Stimuli in the Subject Condition

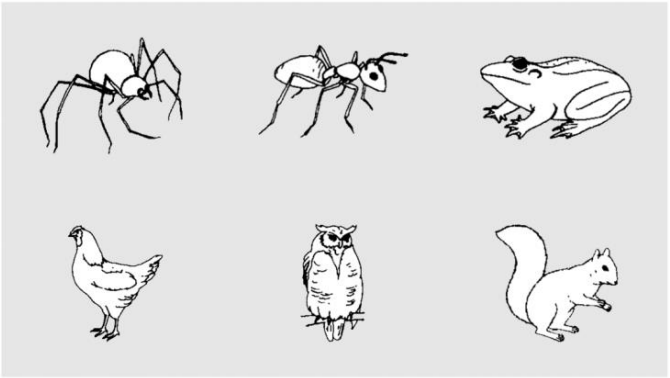
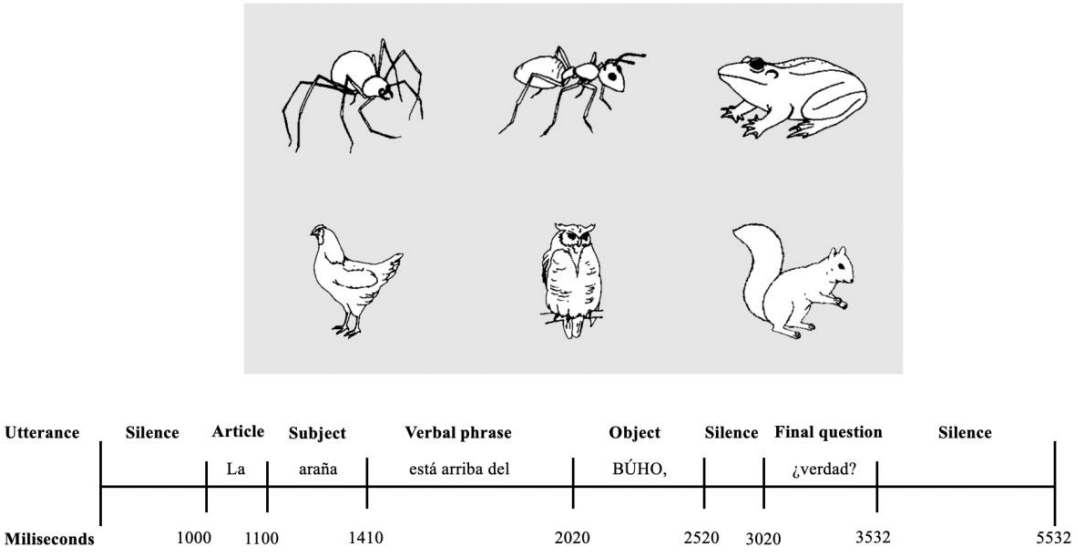


Figure 15

Timestamps of Sentence Sections of Auditory Stimuli in the Prepositional Predicative Condition



Before the analysis, data was first processed, where an interpolation of 150 ms was applied to reconstruct data loss by blinking. In addition, to eliminate noise due to lack of gaze precision, a Gaussian filter was applied.

Then, the proportion of data per trial was calculated, and trials with less than 20% of data on the window of analysis were eliminated, following other studies where a similar criterion was used (10% in the case of Angulo-Chavira & Arias-Trejo, 2021 and Arias-Trejo et al., 2022; 17% in the case of Quam and Swingley, 2014; and 20% in the case of Borovsky et al., 2016). In this case, only the responses to the experimental stimuli were analysed. A total of 1278 from 1512 trials remained in the SC, which accounted for 84.52% of the trials; while in the PPC, a total of 1182 from 1338 trials remained in the SC, which accounted for 88.34% of the trials.

After this procedure, only participants with less than 50% track loss were considered for the analysis, as done before with eye-tracking studies with preschool children (Höhle et al., 2016;

Reuter et al., 2023). For this purpose, the percentage of gazes per participant in each condition was measured. Two participants from the SC were eliminated (*n*= 19), which accounted for 2.35% of data loss; in the PPC, one participant was eliminated (*n*= 18), which accounted for 2.54% of data loss. At the end, 82.54% of the trials remained in the SC and 86.10% of the trials remained in the PPC.

The first analysis consisted of a cluster-based nonparametric analysis (Maris & Oostenveld, 2007). This analysis improves the sensitivity of the statistical tests by solving the multiple comparisons problem in data with multiple time stamps such as eye-tracking data, in which each pair of time stamps needs to be compared and hence the statistical tests may have less sensitivity. In this case, *t*-tests are made to compare two variables, but the result may have to do with the order of the stimuli. For this purpose, the cluster-based nonparametric analysis creates multiple orders of the data and then makes the *t*-tests to every order to ensure the results observed had nothing to do with the way data was arranged.

This method has been used in previous eye-tracking experiments with toddlers (Von Holzen & Mani, 2012; Angulo-Chavira & Arias-Trejo, 2021; Arias-Trejo, et al., 2022; Chow, et al., 2022). It formulates a null hypothesis, which states the equal probability of distribution in the experimental conditions, to prove that the nonparametric test controls the rate of falsely rejecting this null hypothesis.

Each condition, SC and PPC, was analysed separately; looks to the distractors were averaged to eliminate individual effects (Höhle et al., 2016). A difference in looks is expected depending on the emphasis on the subject and the prepositional predicative in both conditions.

The data was divided in bins of 100 ms for the analysis because of the different resolutions of the participants' cameras, which might have slowed the signal (Reuter et al., 2023). This means

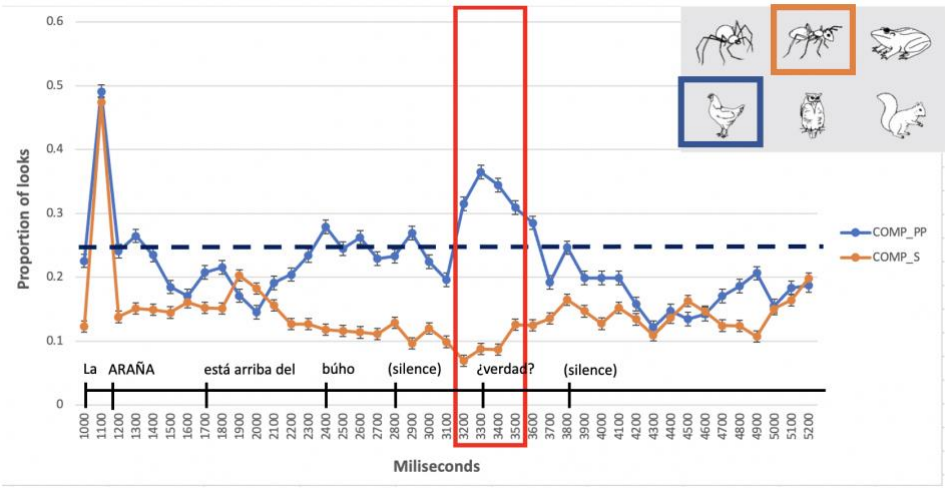
that, in the window of analysis of each trial, data was divided every 100 ms in order to make comparisons between trials. To compare the differences between looks to the subject competitor and the prepositional predicative competitor, *t*-tests were used. For each condition, clusters were created, were *t*-values were equal or greater than a critical value of 0.05 (used for two-tailed hypothesis). Then, the differences between looks to the subject competitor and the prepositional predicative competitor were assigned with positive and negative values randomly, creating 1000 iterations, and *t*-tests were applied for each iteration. This created a permuted distribution. Finally, clusters were identified; and these clusters were considered significant if their *t*-values were equal or greater than the *t*-value of the observed data. A *p* value was then obtained by dividing the number of significant clusters between the total number of iterations, which were 1000.

2.14.3. Results of eye tracking data

In the SC, the prepositional predicative competitor had a greater proportion of looks than the subject competitor from 3200 to 3500 ms (*tclust*= 10.484; *tmax*= 3.690; *p*= 0.001), as seen in Figure 16. Note that the high proportion of looks to both competitors from 1000 to 1100 ms was due to a sampling error with one participant.

Figure 16

Looks in the Subject Condition in the Analysis Window

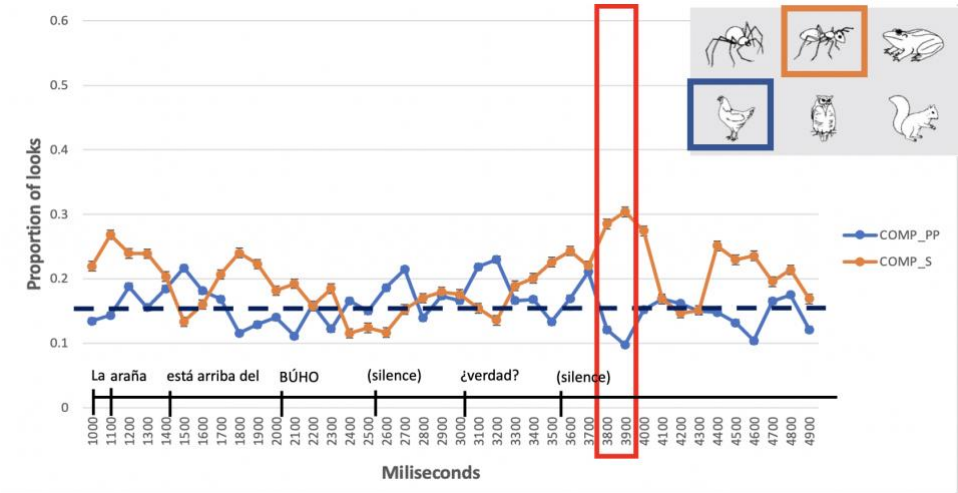


Note. Example of auditory stimulus heard: "The SPIDER is above the owl, right?". COMP_PP refers to the prepositional predicative competitor; COMP_S, to the subject competitor. The Y axis shows the proportion of looks and the X axis shows the time in ms.

In comparison, in the PPC, the subject competitor had a greater proportion of looks than the prepositional predicative competitor from 3800 to 3900 ms (*tclust*= 4.803; *tmax*= 2.890; *p*=0.0005), as seen in Figure 17:

Figure 17

Looks in the Prepositional Predicative Condition in the Analysis Window



Note. Example of auditory stimulus heard: "The spider is above the OWL, right?". COMP_PP refers to the prepositional predicative competitor; COMP_S, to the subject competitor. The Y axis shows the proportion of looks and the X axis shows the time in ms.

In both cases, there was a greater proportion of looks to the prepositional predicative or subject competitor after the prepositional predicative of the sentence was heard; however, in the SC, this occurs earlier, before the final silence.

2.14.4. Correlation of the results of Experiment 1 with the results of the Flanker Task

The same 36 participants, as in Experiment 1, were considered for this analysis (SC, n= 18; PPC, n= 17). In the case of analysis criteria, as in Experiment 1, Spearman's correlation coefficient was used, this time to find correlations between the proportion of looks to the subject, the proportion of looks to the prepositional predicative, and interference effect.

The results showed that, in the SC, no significant correlation was found between the interference effect and the proportion of looks to the subject (r (17) = -0.102, p= 0.687), or

between the interference effect and the proportion of looks to the prepositional predicative (r (17) = 0.193, p= 0.443). In the same line, in the PPC, no significant correlations were found between the interference effect and the proportion of looks to the subject (r (16) = -0.358, p= 0.158) or between the interference effect and the proportion of looks to the prepositional predicative (r (16) = 0.346, p= 0.174).

2.15. Discussion of Experiment 2

Experiment 2 aimed to evaluate the comprehension of prosodic focus in utterances by analysing participants' gaze responses, eliminating the cognitive demands associated with verbal responses observed in Experiment 1. Unlike the previous experiment, the goal was to assess prosodic focus comprehension while utterances were heard.

In this second experiment, the images' position, as shown in the visual stimuli, was incorrectly described in the audio stimuli. The objective was to determine if participants would direct their gaze to the image above or below the emphasized element. If the prosodic focus was on the subject (SC), the correct response would be to look at the image above the prepositional predicative (the correct subject). Conversely, if the prosodic focus was on the prepositional predicative (PPC), the correct response would be to look at the image below the subject (the correct prepositional predicative). The hypothesis was that there would be a greater number of correct gaze responses for PPC compared to SC; and that there would be a greater number of correct responses to the SC than in Experiment 1, aligning with previous research in languages like German (Höhle et al., 2016).

Contrary to expectations, the results revealed that pre-schoolers tended to look at the correct subject in PPC and at the correct prepositional predicative in SC. An interpretation suggests that the emphasized element in the audio stimuli was perceived as the topic of the sentence (that is,

the main theme of the utterance), likely due to its salience and the ease of retention in pre-schoolers 's working memory (as established by Szendröi, 2004). Because the emphasized element was the theme of the utterance, the looks directed to the other element that was below or above this element.

Notably, in SC, looks directed towards the prepositional predicative competitor occurred earlier, suggesting that emphasis in the subject served as an early cue, influencing participants' interpretation of the utterance. In contrast, PPC lacked such cues, as the emphasis occurred at the end, delaying the initiation of alternative looks. In consequence, pre-schoolers of 4 and 5 years of age perceived changes in intonation; and these changes helped them to interpret the utterances they heard in a different manner depending on where the emphasis was in the utterance.

These results differ from the ones of Höhle et al. (2016) of German children of the same age. In the research of Höhle et al. (2016), the eye-tracking experiment showed a greater proportion of looks to the subject in the pre-subject condition and a greater proportion of looks to the object in the post-subject condition.

Nevertheless, it is important to note that in Höhle's et al. (2016) study the sentences used were different than the ones of this dissertation. The focus-sensitive particle "only" was used, guiding the position of the focused element (before or after the subject) in sentences like "Only the ELEPHANT has a kite" or "The elephant only has a KITE". In the case of the present dissertation, however, sentences without "only" are used because a more natural conversation was preferred, where focus is used to correct previous information (Szendröi et al., 2018). Because of this, utterances that elicited contrastive focus, plainly describing the images, were used.

Even so, in concordance to the study of Höhle at al. (2016), in this dissertation, though eye-tracking, a difference of interpretation of prosodic focus between conditions was found, something that did not happen (at least, not with statistical significance) when verbal responses were given. This shows that there may be other cognitive processes involved when answering an experimental task verbally, and that the eye-tracker may be a reliable method to assess online comprehension.

Moreover, no correlations were identified between inhibitory control and the proportion of looks to subjects or prepositional predicatives in both conditions. This absence of correlation may be attributed to the nature of online experiments, such as eye-tracking, which assess sentence comprehension as it occurs, potentially involving fewer cognitive processes like inhibitory control (Szendröi, 2004; Höhle et al., 2016).

3. General Discussion

The present study aimed to assess the comprehension of prosodic focus in Mexican Spanish-speaking pre-schoolers of 4 and 5 years of age in utterances with emphasis on the subject (SC) or the prepositional predicative (PPC). The primary aim was to examine how infants in this population understand changes in prosody, specifically focusing on how prosodic emphasis is perceived in distinct positions within Spanish utterances. The results indicated that preschool Spanish-speaking children could distinguish prosodic emphasis at the beginning and end of utterances; however, this was without a clear focus interpretation.

Previous research in languages like English, Danish, and German, where focus is mainly marked using prosody, demonstrated that children as young as 3 can use it to interpret utterances (Höhle et al., 2016; Szendröi et al., 2018). These languages rely on changes in intonation due to their fixed word order (that is, the sentences follow a Subject, Verb and Object pattern, which

cannot be changed). In contrast, languages like Spanish, with free word order, use not only prosody, but also syntactic movement to mark focus (Gutiérrez-Bravo, 2008). Thus, understanding how pre-schoolers develop prosodic focus comprehension is crucial for this language, as it is important to investigate how much they rely on prosodic cues to mark focus or if syntactic movement is a stronger and necessary cue.

Despite challenges in assessing prosodic focus in past research, possibly due to issues with the experimental tasks (Ito, 2018), recent suggestions emphasize language differences as contributing factors, especially when prosody is not the sole resource for marking focus (Surányi & Pintér, 2020; 2022). This dissertation, as an exploratory study, fills a gap, as there are no known studies to my knowledge on prosodic focus comprehension in Spanish preschool children.

A first aim was to assess verbal responses to compare results with previous investigations in other languages and, furthermore, to examine if acoustic analysis is a method that helps to better understand the nature of the responses given. In Experiment 1, participants corrected statements with emphasis in the subject or prepositional predicative. It is important to point out that, despite assessing production in this experiment, the prosodic prominence seen in participants' responses depended on where they perceived the emphasis in the auditory stimuli (subject or prepositional predicative).

The hypothesis was that pre-schoolers will tend to correct the prepositional predicative more frequently than the subject in both conditions. Results showed that, no matter if emphasis lied on the subject or prepositional predicative of the auditory stimuli, most verbal responses corrected the prepositional predicative of the sentence, confirming this hypothesis. Additionally, there were no differences in the tone patterns between conditions.

These results aligned with French, another language with free word order using syntactic movement to mark focus (Szendröi et al., 2018). It seems that languages that do not use prosody alone for focus may show slower prosodic focus acquisition, as suggested by recent research in Hungarian (Surányi & Pintér, 2020; 2022), a language that also uses syntactic to mark focus, but using verbal particles.

In their research, Surányi & Pintér (2020; 2022) showed that monolingual Hungarian-speaking pre-schoolers of 4 years of age also had a greater number of errors that children of these ages who speak English, German and even French, as they had difficulties interpreting focus at the beginning, as it usually falls at the end. These errors diminished with age, at around 6 years. The authors also used a study based on Szendröi et al. (2018) as this present investigation did, which further shows how prosodic focus acquisition may also be language dependent. In consequence, Spanish-speaking pre-schoolers may have greater difficulties to comprehend focus with prosody alone at preschool years.

Even so, at least some pre-schoolers corrected the subject in both conditions, most of them being in the PPC and not the SC in the present research. According to Szendröi (2004), pre-schoolers that do interpret focus at a position other than the end of the utterance (in this case, in the subject), could possibly have more developed cognitive abilities. As focus usually lies at the end of the sentences, working memory or inhibitory control may play a role when answering to utterances that have focus on another position (Szendröi, 2004; Höhle, 2016). In this sense, it seems that pre-schoolers that corrected the subject more often had a less developed inhibitory control, according to the correlations found with the Flanker Task. The interpretation of these results could be better understood looking at the results of Experiment 2.

The goal of Experiment 2 was to assess online comprehension of heard utterances, acknowledging that verbal responses might involve additional cognitive abilities. Unexpectedly, the results revealed a greater proportion of looks to the prepositional predicative competitor in the SC and to the subject in the PPC, the element that would be corrected. This pattern, contrary to our expectations (more looks directed towards the subject competitor in the SC and towards the prepositional predicative in the PPC), suggests that pre-schoolers can distinguish emphasis but interpret it differently than focus.

It is plausible that the emphasized element is perceived as the topic of the sentence, making the other element appear as the one needing correction. This interpretation aligns with the idea that the emphasized element, being more salient, remains in children's memory longer, effectively becoming the theme or topic of the utterance. Consequently, pre-schoolers direct their gaze to the correct prepositional predicative related to the emphasized subject in the SC, and to the correct subject related to the emphasized prepositional predicative in the PPC.

Given that cognitive processes sometimes interfere with verbal responses in experimental tasks (Höhle et al., 2016), Experiment 2 employed an eye-tracker to gain a better understanding of prosodic focus interpretation as participants listened to auditory stimuli. Comparing these results with those of Experiment 1, where pre-schoolers correcting the subject more often were associated with less developed inhibitory control, it's plausible that some children struggled to inhibit a standard response where the prepositional predicative was considered the topic of the sentence, leading them to correct the subject. However, this interpretation should be approached with caution, given the marginally significant correlation and the potential for varied interpretations of emphasis among pre-schoolers.

Recent research in Hungarian (Pintér & Surányi, 2023), which, like Spanish, uses syntactic movement and prosody to mark focus, has highlighted the importance of context. Pre-schoolers may face challenges in identifying the alternatives to which contrastive focus refers. Consequently, tasks like the ones used by Szendröi et al. (2018), and hence like the one used in this dissertation, may be less clear to some pre-schoolers, needing more explicit questions (e.g., "Who is above/below the ___?") to interpret emphasis as focus. This could explain why pre-schoolers in this dissertation interpreted emphasis differently than focus.

Additionally, variations in how pre-schoolers of this population interpreted emphasis might be attributed to duration and intensity as more reliable forms of marking focus, as observed in adults from central Mexico (De-la-Mota, Martín-Butragueño, & Prieto, 2010). Another possibility is that pre-schoolers use different acoustic cues to mark prosodic focus than adults, such as pauses between the focused element and the rest of the utterance or different tone patterns (Villalobos-Pedrosa, 2021). Results from the Pilot Test of Experiment 1, in the acoustic analysis, showed similarities to those in Experiment 2. This may have happened because stimuli in the Pilot Test included a pause between the subject and the rest of the utterance in the SC, a cue that pre-schoolers use to mark prosodic focus (Villalobos-Pedrosa, 2021), which was later eliminated in the final experimental task. Future studies could explore this theory and compare it with the results of an eye-tracking study. Nevertheless, this investigation showed that pre-schoolers' do rely on prosodic information to interpret sentences.

4. Conclusions

As previously found in research in other languages (Szendröi, 2004; Höhle et al., 2016; Szendröi et al., 2018), this investigation showed that Mexican preschool children are able to use

prosodic emphasis alone as a cue to interpret sentences in Spanish, a claim supported by eye-tracker data.

This investigation, however, has some limitations. Firstly, in Experiment 1, a more detailed acoustic analysis is warranted in future research, specifically focusing on more specific criteria for pitch classification based on the position of the intonational curve's peak inside the tonic syllable, which was not considered in this initial analysis. Moreover, the sentences used may not have elicited an interpretation of focus in the emphasized element of the utterance, and additional context might be necessary to evoke a focus interpretation, as suggested by Pintér and Surányi (2023). Additionally, other types of utterances, such as transitive sentences, could be considered in the future for more precise result comparisons with those of Szendröi et al. (2018). Finally, pauses could be considered as cues to mark focus, which pre-schoolers find more reliable (as observed in Villalobos-Pedrosa, 2021), an aspect that was not taken into account in the final stimuli used in this experiment. Duration and intensity could also be regarded as new parameters to mark focus, given their preference among the population of the central Mexico (De-la-Mota, Martín-Butragueño, & Prieto, 2010).

Despite these limitations, the fact that, in Experiment 2, pre-schoolers were able to direct their gaze to the correct element earlier in the SC, where prosodic focus is presented at the beginning of the sentence, and later in the PPC, where prosodic focus was marked at the end, demonstrates that pre-schoolers can interpret syntactically identical sentences differently depending on the position of prosodic focus.

Future research may consider a more detailed acoustic analysis, test focus on different types of utterances, and consider other cues to mark focus. It would also be interesting to compare the obtained results in this investigation to the ones of adult speakers. The Pilot Test of Experiment 2

(Appendix H), which exclusively used adult speakers, revealed a difference in looks to the subject and prepositional predicative competitors only in the SC, with a greater proportion of looks to the prepositional predicative, similar to what we observed in pre-schoolers. This could be further examined with diverse types of focused sentences.

In conclusion, the current study provides valuable insights into how pre-schoolers interpret prosodic focus in a language with free word order. It is a pioneering approach that can be explored with nuanced research designs and in consideration of language-specific cues in future research.

5. References

Adell, F., Aguilar, L., Corrales-Astorgano, M., y Escudero-Mancebo, D. (*s.f.*) Proceso de innovación educativa en educación especial: enseñanza de la prosodia con fines comunicativos con el apoyo de un videojuego educativo. In *Proceedings of the I Congreso Internacional en Humanidades Digitales, Valladolid, Spain* (pp. 17-19).

Anderson, M. C., & Weaver, C. (2009). Inhibitory Control over Action and Memory. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (pp. 153-163). Elsevier.

Angulo-Chavira, A. Q., & Arias-Trejo, N. (2021). Mediated semantic priming interference in toddlers as seen through pupil dynamics. *Journal of Experimental Child Psychology*, *208*, 105146. https://doi.org/10.1016/j.jecp.2021.105146

Arias-Trejo, N., Angulo-Chavira, A. Q., Avila-Varela, D. S., Chua-Rodriguez, F., & Mani, N. (2022). Developmental changes in phonological and semantic priming effects in Spanish-speaking toddlers. *Developmental Psychology*, *58*(2), 236. https://doi.org/10.1037/dev0001290

Authier, J. M., & Haegeman, L. (2018). The Syntax of Mirative Focus Fronting: Evidence from French. In D. L. Arteaga (Ed.), *Contribution of Romance Languages To Current Linguistic Theory* (pp. 39-64). Springer.

Baddeley, A. D. (1992). Working Memory. *Science*, *255*(5044), 556–559. http://doi.org/10.1126/science.1736359

_____ (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, (63), 1-29. http://doi.org/10.1146/annurev-psych-120710-100422

Baeza-Álvarez, R. L., & Rodríguez-Maldonado, D. D. (2011). Comprensión de la prosodia en el lenguaje verbal de niños escolares. *Revista de Educación y Desarrollo*, (17), 5–14.

Beckman, M., M. Díaz-Campos, J. Tevis Mc Gory, & T. Morgan. (2002). Intonation across Spanish, in the tones and break indices framework. *Probus* 14, 9-36. http://doi.org/10.1515/prbs.2002.008

Best, J. R. & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Dev*, 81 (6), 1641-1660. http://doi.org/10.1111/j.1467-8624.2010.01499.x

Boersma, P., & Weenink, D. (2006). Praat: Doing Phonetics by Computer. Available at: http://www.praat.org/

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: Category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental science*, *19*(6), 918-932. https://doi.org/10.1111/desc.12343

Brickman, A. M., & Stern, Y. (2009). Aging and memory in humans. In P. R. Hof & C. V. Mobbs (Eds.), *Handbook of the neuroscience of aging* (pp. 243-248). Elsevier Academic Press.

Chafe, W. (1974). Language and consciousness. *Language*, 1974, 50, 111-133. https://doi.org/10.2307/412014

Chen, A. (2010). Is there really an asymmetry in the acquisition of the focus-to-accentuation mapping? *Lingua*, *120*(8), 1926–1939. http://doi.org/10.1016/j.lingua.2010.02.012

Chow, J., Angulo-Chavira, A. Q., Spangenberg, M., Hentrup, L., & Plunkett, K. (2022). Bottom-up processes dominate early word recognition in toddlers. *Cognition*, *228*, 105214. https://doi.org/10.1016/j.cognition.2022.105214

Cinque, G. (1993) A Null Theory of Phrase and Compound Stress. *Linguistic Inquiry*, *24*(2), 239-297. https://www.jstor.org/stable/4178812

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (2008). Variation in working memory: An introduction. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 3-17). Oxford University Press.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language*, *12*(3), 643–661. http://doi.org/10.1017/S030500090000670X

Cutler, A., & Swinney, D. A. (1987). Asymmetry from an intonation-universal based account perspective. Prosody and the development of comprehension. *Journal of Child Language*, *14*, 145–167.

De Carvalho, A., Dautriche, I., Millotte, S., & Christophe, A. (2018). Early perception of phrasal prosody and its role in syntactic and lexical acquisition. In P. Prieto & N. Esteve-Gibert. (Eds.), *The Development of Prosody in First Language Acquisition.* (pp. 245–270). John Benjamins Publishing Company. http://doi.org/10.1075/tilar.23.02car

De Carvalho, A., He, A. X., Lidz, J., & Christophe, A. (2019). Prosody and Function Words Cue the Acquisition of Word Meanings in 18-Month-Old Infants. *Psychological Science*, 30(3). http://doi.org/10.1177/0956797618814131

De-la-Mota, C., Martín Butragueño, P., & Prieto, P. (2010). Mexican Spanish intonation. *Transcription of intonation of the Spanish language*, 319-350.

Diamond, A. (2013). Executive Functions. *Annu Rev Psychol*, 64, 135-168. http://doi.org/10.1146/annurev-psych-113011-143750

Dik, S. C. (1997). *The theory of functional grammar. Part 1: The structure of the clause*. De Gruyter Mouton.

Dorta-Luis, J. (2008). La focalización prosódica: funcionalidad en los niveles lingüístico y pragmático. *Estudios de fonética experimental*, XVII, 106-138.

Face, T. (2002). El foco y la altura tonal en el español. *Boletín de Lingüística, 14*(17), 30-52.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Féry, C. (2013). Focus as prosodic alignment. *Natural Language & Linguistic Theory, 31*(3), 683–734. https://doi.org/10.1007/s11049-013-9195-7

Fifer, W.P., & Moon, C.M. (1994). The role of mother's voice in the organization of brain function in the newborn. *Acta Paediatr*, 397, 89–93. https://doi.org/10.1111/j.1651-2227.1994.tb13270.x

Furrow, D. (1984). Young children's use of prosody. *Journal of Child Language,* 203–213. http://doi.org/10.1017/S0305000900005663

Garf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy,* 18(5), 797-824. http://doi.org/10.1111/infa.12006

Grassmann, S., & Tomasello, M. (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language*, 34(3), pp. 677-687. http://doi.org/10.1017/S0305000907008021

Grassmann, S., & Tomasello, M. (2010). Prosodic stress on a word directs 24-month-olds' attention to a contextually new referent. *Journal of Pragmatics*, 42(11), pp. 3098-3105. https://doi.org/10.1016/j.pragma.2010.04.019.

Goswami, U. (2008). *Cognitive Development: The Learning Brain*. Psychology Press.

Gualmini, A., Maciukaite, S., & Crain, S. (2002). Children's Insensitivity to Contrastive Stress in Sentences with ONLY. In *Proceedings of the 26th Annual PCL*, *8*(1), 1–14.

Gutiérrez-Bravo, R. (2008). La identificación de los tópicos y los focos. *Nueva Revista de Filología Hispánica*, 56(2), 363-401. https://www.jstor.org/stable/40300721

Hendriks, P., & Koster, C. (2010). Comprehension asymmetries in language acquisition. *Lingua, 120*(8), 1887–1897. http://doi.org/10.1016/j.lingua.2010.02.002

Höhle, B., Fritzsche, T., & Müller, A. (2016). Children's Comprehension of Sentences with Focus Particles and the Role of Cognitive Control: An Eye Tracking Study with German-Learning 4-Year-Olds. *PloS One*, *11*(3), e0149870. http://doi.org/10.1371/journal.pone.0149870

Hornby, P.A., & Hass, W.A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research,* 13, 359–399. http://doi.org/10.1044/jshr.1302.395

Hornby, P.A. (1971). Surface structure and the topic-comment distinction: a developmental study. *Child Development*, 42(6), 1975–1988. https://doi.org/10.2307/1127600

Hualde, J. I. (2013). *Los sonidos del español: Spanish language edition*. Cambridge University Press.

Hualde, J. & Prieto, P. (2015). Intonational variation in Spanish: European and american varieties. En S. Frota y P. Prieto (Eds.), *Intonational Variation in Romance* (pp. 350-391). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199685332.003.0010

Ito, K. (2018). Gradual development of focus prosody and affect prosody comprehension. In P. Prieto & N. Esteve-Gibert. (Eds.), *The Development of Prosody in First Language Acquisition.* (pp. 245–270). John Benjamins Publishing Company. http://doi.org/10.1075/tilar.23.13ito

Jackson-Maldonado, D., Thal, D., Marchman, V., Newton, T., Fenson, L, & Conboy, B. (2003). *MacArthur Inventarios del Desarrollo de Habilidades Comunicativas. User´s Guide and Technical Manual*. Brookes Publishing.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of memory and language*, *44*(4), pp. 548-567. https://doi.org/10.1006/jmla.2000.2755

Jun, Sun-Ah (2005). Prosodic typology. En Sun-Ah Jun (Ed.), *Prosodic typology*: *The phonology of intonation and phrasing* (pp. 430–458). Oxford University Press. https://doi.org/10.1093/acprof:oso/978019924 9633.001.0001

Kisilevsky, B.S., et al. (2003). Effects of experience on feral voice recognition. *American Psychological Society, 14*(3), 220-224. http://doi.org/10.1111/1467-9280.02435

Kisilevsky, B.S., et al. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32(1), 59–71. http://doi.org/10.1016/j.infbeh.2008.10.002

Kuhl, et al. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society*, 363, 979-1000. http://doi.org/10.1098/rstb.2007.2154

Kumin, L. (2014). *Síndrome de Down: habilidades tempranas de comunicación. Una guía para padres y profesionales.* Fundación Iberoamericana Down21.

Ladd, R. D. (2008). *Intonational Phonology* (*2ª* ed.). Cambridge University Press.

Lahey, M. (1974). Use of Prosody and Syntactic Markers in Children's Comprehension of Spoken Sentences. *Journal of Speech and Hearing Research*, 17, 656-668. http://doi.org/10.1044/jshr.1704.656

Lamers, K. y Hall, L. J. (2003). The Response of Children with Autism to Preferred Prosody During Instruction. *Focus on Autism and Other Developmental Disabilities*, *18*(2), 95–104.

Liu, Q., Zhu, X., Ziegler, A., & Shi, J. (2015). The effects of inhibitory control training for preschoolers on reasoning ability and neural activity. *Scientific reports,* 5 (14200), 1-10. http://doi.org/10.1038/srep14200

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant-directed and adult-directed speech. *Language Learning and Development*, 7(3), 209-225. http://doi.org/10.1080/15475441.2011.579839

MacWhinney, B., & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior*, *17*(5), 539–558. http://doi.org/10.1016/S0022-5371(78)90326-2

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Martín-Butragueño, P. (2014). *Fonología variable del español de México. Volumen 1: procesos segmentales, México.* El Colegio de México.

Martín-Butragueño, P. (2019). *Fonología variable del español de México. Volumen II: prosodia enunciativa. Tomo I.* El Colegio de México.

Martin, N. A. (2013). *Receptive One-Word Picture Vocabulary Test-4. Spanish-Bilingual Edition*. ATP Assessments, a division of Academic Therapy Publications.

McDaniel, D., & Maxfield, T. L. (1992). Principle B and Contrastive Stress. *Language Acquisition*, *2*(4), 337–358. http://doi.org/10.1207/s15327817la0204_4

Minai, U., Jincho, N., Yamane, N., & Mazuka, R. (2012). What hinders child semantic computation: children's universal quantification and the development of cognitive control. *Journal of Child Language*, 39 (5), 919–56. http://doi.org/10.1017/S0305000911000316

Moon, C., Lagercrantz, H., & Kuhl, K.P. (2012). Language experienced affects vowel perception after birth: a two-country study. *Acta Paedriatrica*, *102*(2), 156–160. http://doi.org/10.1111/apa.12098

Mora-Bustos, A. (2010). Construcciones escindidas, foco y entonación. In S. Bogard. (Ed.), *Semántica, pragmática y prosodia. Reflejos en el orden de palabras en español.* (pp. 219-248). Centro de Estudios Lingüísticos y Literarios - El Colegio de México.

Nazi, T., Bertoncini, J., & Mehler, J. (1998). Language Discrimination by Newborns: Toward an Understanding of the Role of Rhythm. *Journal of Experimental Psychology*, *24*(3), 756–766. http://doi.org/10.1037//0096-1523.24.3.756

Neisser, U. (2014). *Cognitive Psychology. Classic Edition*. Psychology Press.

Nespor, M. & Vogel. I. (1994). *La Prosodia*. Visor.

Nigg, J. T. (2000). On Inhibition/Disinhibition in Developmental Psychopathology: Views From Cognitive and Personality Psychology and a Working Inhibition Taxonomy. *Psychological Bulletin*, 126 (2), pp. 220-246. http://doi.org/10.1037//0033-2909.126.2.220

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, and Behavioral Neuroscience*, (3), pp. 263–81. http://doi.org/10.3758/CABN.5.3.263

Ojeda López, A. I. (2000). El desarrollo sintáctico en la adquisición de la primera lengua: análisis de la etapa telegráfica de una sujeto monolingüe de español. *Revista Española de Lingüística*, 31 (2), pp. 413-430.

Paller, K. A. (2009). Memory Consolidation: Systems. In L. R. Squire (Ed.), *Encyclopedia of neuroscience*. (pp. 741-749). Elsevier.

Pamies Bertrán, A., A. M. Fernández Planas, E. Martínez Celdrán, A. Ortega Escandell, & Amorós Céspedes M. (2001). Umbrales tonales en español peninsular. In M. Cuenca Villain (Ed.), *Actas del II Congreso Nacional de Fonética Experimental* (pp. 272-278). Universidad de Sevilla.

Paterson, K. B., Liversedge, S. P., Rowland, C., & Filik, R. (2003). Children's comprehension of sentences with focus particles. In *Cognition, 89*(3), 263–294. http://doi.org/10.1016/S0010-0277(03)00126-4

Paterson, K. B., Liversedge, S. P., White, D., Filik, R., & Jaz, K. (2006). Children's Interpretation of Ambiguous Focus in Sentences With "Only". In *Language Acquisition, 13*(3), 253–284. http://doi.org/10.1207/s15327817la1303

Perrone-Bertolotti, M., Dohen, M., Loevenbruck, H., Sato, M., Pichat, C., & Baciu, M. (2013). Neural correlates of the perception of contrastive informative focus in French: a functional magnetic resonance imaging study. *Human brain mapping, 34*(10), 2574–2591. http://doi.org/10.1002/hbm.22090

Pintér, L., & Surányi, B. (2023). Preschoolers' comprehension of exhaustive focus: The role of contextualization. *First Language, 43*(1), 58-90. https://doi.org/10.1177/01427237221113962

Pynte, J. (1998). The Role of Prosody in Semantic Interpretation. *Music Perception: An Interdisciplinary Journal, 16*(1), 79–97. http://doi.org/10.2307/40285779

Quam, C., & Swingley, D. (2014). Processing of lexical stress cues by young children. *Journal of experimental child psychology*, *123*, 73-89. https://doi.org/10.1016/j.jecp.2014.01.010

Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Editorial Gredos.

Rains, D. (2004). Lenguaje. In *Principios de Neuropsicología Humana*. (125–151). Mc Graw Hill.

Reinhart, T. (1999). The processing cost of reference-set computation: guess patterns in acquisition. In *OTS Working Papers in Linguistics*.

Reuter, T., Mazzei, C., Lew-Williams, C., & Emberson, L. (2023). Infants' lexical comprehension and lexical anticipation abilities are closely linked in early language development. *Infancy*, *28*(3), 532-549. https://doi.org/10.1111/infa.12534

Reyzábal, M. I., & Santiuste, V. (2005). Modelos del lenguaje y tecnología del habla. *Revista Educación XX1*, *8*(3), 127–134. https://doi.org/10.5944/educxx1.8.0.346

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004) Development of attentional networks in childhood. *Neuropsychologia*, 42 (8), 1029-40. http://doi.org/10.1016/j.neuropsychologia.2003.12.012

Shabidullah, S., & Hepper, P. G. (1994). Frequency discrimination by fetus. *Early Human Development*, *36*(1), 13-26. http://doi.org/10.1016/0378-3782(94)90029-9

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Solan, L. (1980). Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research*, (23), 688–698. http://doi.org/10.1044/jshr.2303.688

Surányi, B., & Pintér, L. (2022). Children's comprehension of prosodically marked focus in Hungarian: How mandatory syntactic focus-marking affects the trajectory of acquisition. *Journal of Child Language*, *49*(4), 824-838. https://doi.org/10.1017/S0305000921000313

Surányi, B., & Pintér, L. (2020). The Acquisition of Prosodic Focus-Identification: The Role of Variation in Focus-Marking. In M. M. Brown & A. Kohut (Eds.), *Proceedings of the 44th Boston University Conference on Language Development* (pp. 648-661). Cascadilla Press.

Szendröi, K. (2004). Acquisition evidence for an interface theory of focus. In J. Van Kampen & S. Baauw (Eds.), *Proceedings of GALA 2003*.

Szendröi, K., Bernard, C., Berger, F., Gervain, J., & Höhle, B. (2018). Acquisition of informative focus marking by English, French, and German three-, four-, five- and six-year-olds. *Journal of Child Language, 45*(1), 219–241. http://doi.org/10.1017/S0305000917000071

Tonkova-Yampol'skaya, R. (1973) Development of speech intonation in infants during the first two years of life. In C. A. Ferguson y D. I. Slobin (ed.), *Studies of child language development.* Holt, Rinehart & Winston.

Vallduví, E., & Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics, (*34), 459–519. http://doi.org/10.1515/ling.1996.34.3.459

Vanrell, M. d. M., Stella, A., Gili-Favela, B., &, Prieto, P. (2013). Prosodic manifestations of the Effort Code in Catalan, Italian and Spanish Contrastive Focus, *Journal of the International Phonetic Association*, *43*(2), 195-220. https://www.jstor.org/stable/26351947

Villalobos-Pedroza, L. (2019). La expresión del foco: Marcos léxicos y estrategias prosódicas del habla dirigida a niños en español. In V. Belloro (Ed.). *La interfaz sintaxis-pragmática: Estudios teóricos, descriptivos y experimentales* (pp. 283-310). De Gruyter.

Villalobos-Pedroza, L. (2021). *Prosodia de la expresión temprana del foco informativo y del foco contrastivo en español* [PhD Thesis]. Universidad Nacional Autónoma de México.

Vidal, M. V. E., & Leonetti, M. (2019). Una nota sobre el foco informativo en español. In R. Martínez Ruiz, I. Olza, & O. Loureda Lamas (eds.), *Lengua, cultura, discurso. Estudios ofrecidos al profesor Manuel Casado Velarde* (pp. 203-227). EUNSA.

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569-586. https://doi.org/10.1016/j.jecp.2012.08.001

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. Psychological Science, 24, 2143–2152. http://doi.org/10.1177/0956797613488145

Wieman, L. A. (1976). Stress patterns of early child language. *Journal of Child Language*, *3*(2), 283–286. http://doi.org/10.1017/S0305000900001501

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of pragmatics*, *38*(10), 1559-1579. https://doi.org/10.1016/j.pragma.2005.04.012

Appendix A

Visual and Word Stimuli Validation

The images were selected from Snodgrass and Vanderwart's (1980) database, except for the image of an egg, which was recovered from a free online database (https://www.istockphoto.com). However, a validation was carried out to ensure these images were representative for Mexican children.

For this validation, a total of twenty children (9 females and 11 males) were recruited online using social media accounts from the UNAM Laboratory of Psycholinguistics. Participants' ages ranged from 4;3 to 7;9 years old ($\bar{x}$=5.93, *SD*=1.15).

Procedure of Visual and Word Stimuli Validation

Using the platform FindingFive, a game was created in which the cartoon of a little girl with a paintbrush appeared on screen asking the child to help her make sure her drawings were good. The voice of a female Mexican-Spanish speaker using child-directed speech was used for this matter, and text also appeared on screen in case the parents needed to read the guidelines. The instructions were as follows: "*¡Hola! Gracias por venir. Estoy aprendiendo a dibujar, ¿me puedes decir qué son estos dibujos para saber si lo hice bien?"* ["Hi! Thanks for coming. I am learning how to draw. Can you tell me what these drawings are so that I know if I drew them correctly?"]. Then, the first image appeared on screen, with the instruction "*Dime claro y fuerte, ¿qué es esto?*" ["Tell me loud and clear, what is this?"]. Five seconds passed in which the platform recorded the participants' voices, and then the next image appeared. Before recording each response, the question "*¿Qué es esto?*" ["What is this?"] was presented. At the end, the image of the little girl appeared again saying: "*¡Gracias! Seguiré practicando mi dibujo. ¡Fuiste*

de mucha ayuda! ¡Hasta luego!" ["Thank you! I will continue practicing my drawings. You were of major help! See you later!"].

A total of 18 images were presented, 12 of which were the stimuli considered for the experiment. It was decided that images with at least 60% of naming consistency were going to be used on the experiment. Additionally, only words said by children that were exactly the same as the ones used in the experiment were considered. For this reason, if an image was named by children with a synonym consistently, it was not included in the experiment. For example, most of the children said "*oveja*" instead of "*borrego*" ("sheep"), "*serpiente*" instead of "*víbora*" ("snake), and "*ciervo*" or "*reno*" instead of "*venado*" ("deer"); these synonyms did not consist of only voiced consonants, an important criteria to visualize the intonational curve (Vanrell et al., 2013), and hence these images were changed for others with a better naming consistency.

Results of Visual Stimuli Validation

The results of the validation are presented in Table A1 below:

Table A1

Results from the Image Validation Task

Image	Percentage of children that named the image correctly	Percentage of children that identified the image
Víbora (Snake)	10%	100%
Agua (Water)	45%	100%
***Araña* (Spider)**	**100%**	**100%**
***Ardilla* (Squirrel)**	**90%**	**90%**
Borrego (Sheep)	15%	95%
***Búho* (Owl)**	**95%**	**95%**
Burro (Donkey)	70%	70%
***Gallina* (Hen)**	**70%**	**70%**
***Globo* (Balloon)**	**90%**	**90%**
***Hormiga* (Ant)**	**80%**	**80%**
***Huevo* (Egg)**	**100%**	**100%**
***Libro* (Book)**	**100%**	**100%**
***Llave* (Key)**	**100%**	**100%**
***Mano* (Hand)**	**100%**	**100%**
Mono (Monkey)	60%	85%
***Rana* (Frog)**	**100%**	**100%**
***Uva* (Grape)**	**100%**	**100%**
Venado (Deer)	45%	70%

Note. The image was deemed *correctly identified* if children were able to name it using the expected word or a synonym. However, it was considered as *correctly named* only if the child used the expected word shown in the first column. The selected images are highlighted in bold.

At the end, the images of a spider, an ant, a frog, a hen, an owl, a squirrel, a hand, an egg, grapes, a balloon, a key and a book were used in the experiment because they had 70% of naming consistency or more. This validation thus confirmed that Mexican children knew these words and images.

Appendix B

List of stimuli

Figure B1

First Combination of Practice Images

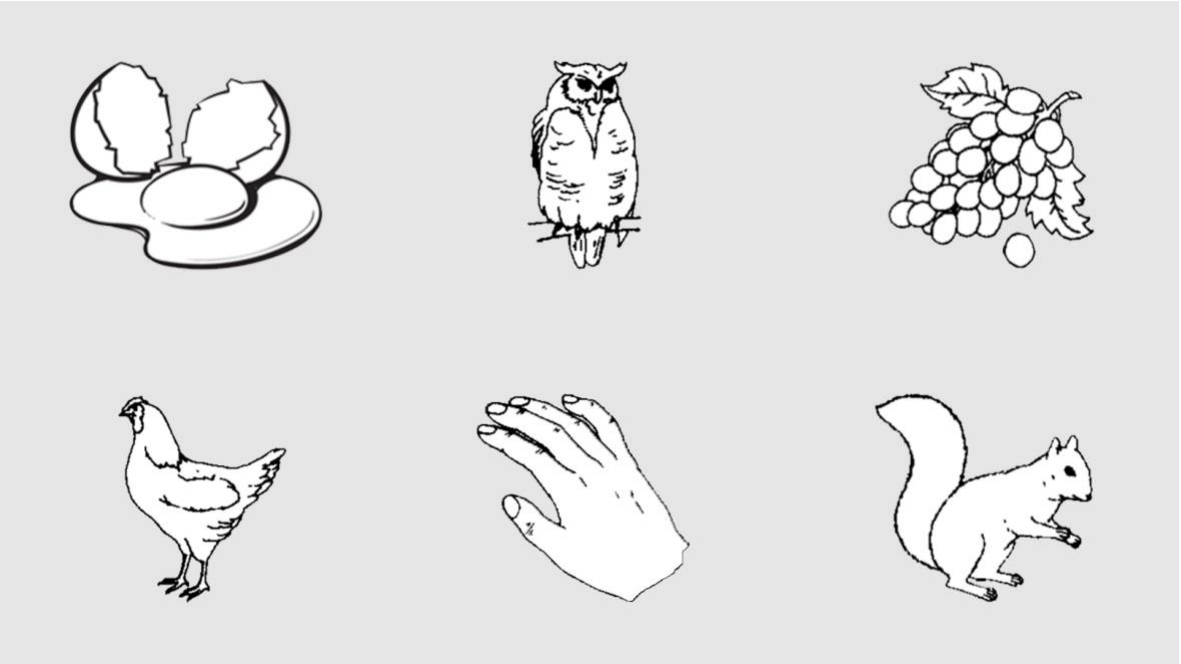


Figure B2

Second Combination of Practice Images

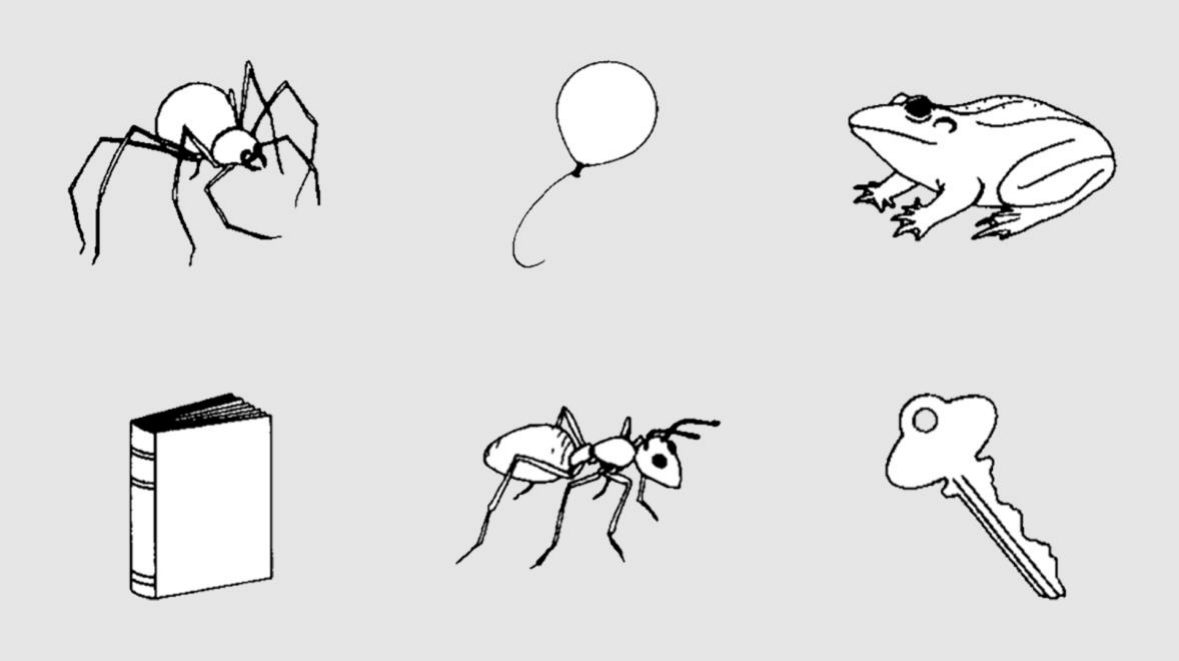


Figure B3

Experimental Images for Table B1

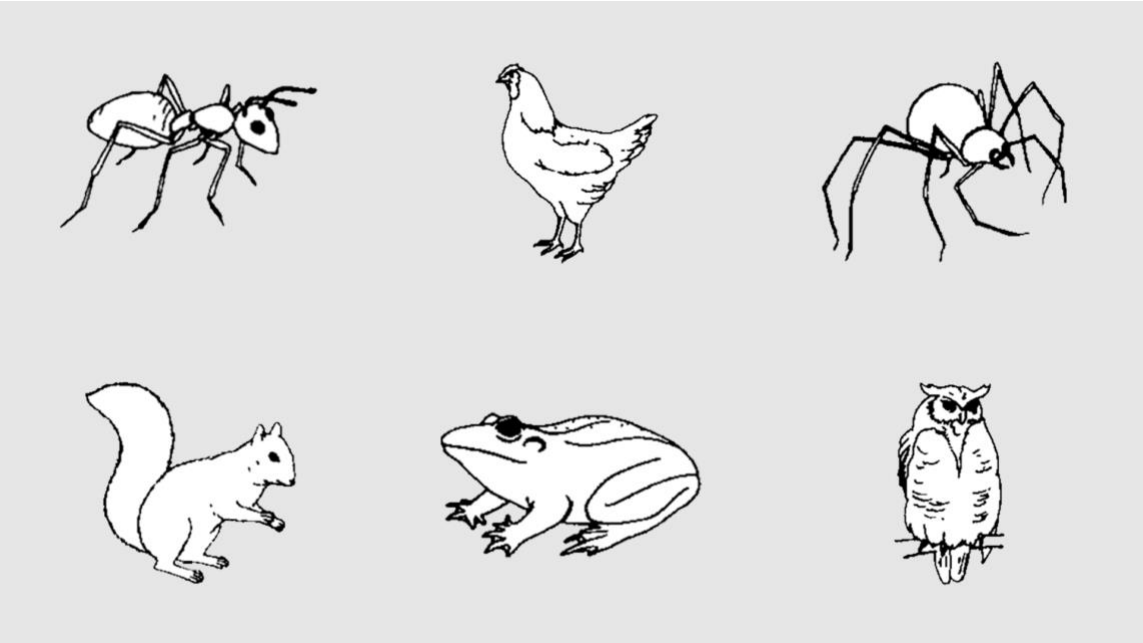


Table B1

Experimental Items for Figure B3

Subject Condition	Prepositional Predicative Condition
1.	*1.*
Q- *La HORMIGA está arriba del búho, ¿verdad?*	Q- *La hormiga está arriba del BÚHO, ¿verdad?*
[Q- The ANT is above the owl, right?]	[Q- The ant is above the OWL, right?]
A- *No, está la ARAÑA.*	A- *No, de la ARDILLA.*
[A- No, the SPIDER is].	[A- No, it is above the SQUIRREL].
2.	*2.*
Q- *La GALLINA está arriba de la ardilla, ¿verdad?*	Q- *La gallina está arriba de la ARDILLA, ¿verdad?*
[Q-The HEN is above the squirrel, right?]	[Q-The hen is above the SQUIRREL, right?]
A- *No, está la HORMIGA.*	A- *No, de la RANA.*
[A- No, the ANT is].	[A- No, it is above the FROG].
3.	*3.*
Q-La ARAÑA está arriba de la rana, ¿verdad?	*Q-La araña está arriba de la RANA, ¿verdad?*
[Q- The SPIDER is above the frog, right?]	[Q- The spider is above the FROG, right?]
A- *No, está la GALLINA.*	A- *No, del BÚHO.*
[A- No, the HEN is].	[No, it is above the OWL].

Note. "Q" refers to the question asked by the experimenter and "A" to the expected response given by pre-schoolers.

Figure B4

Experimental Images for Table B2

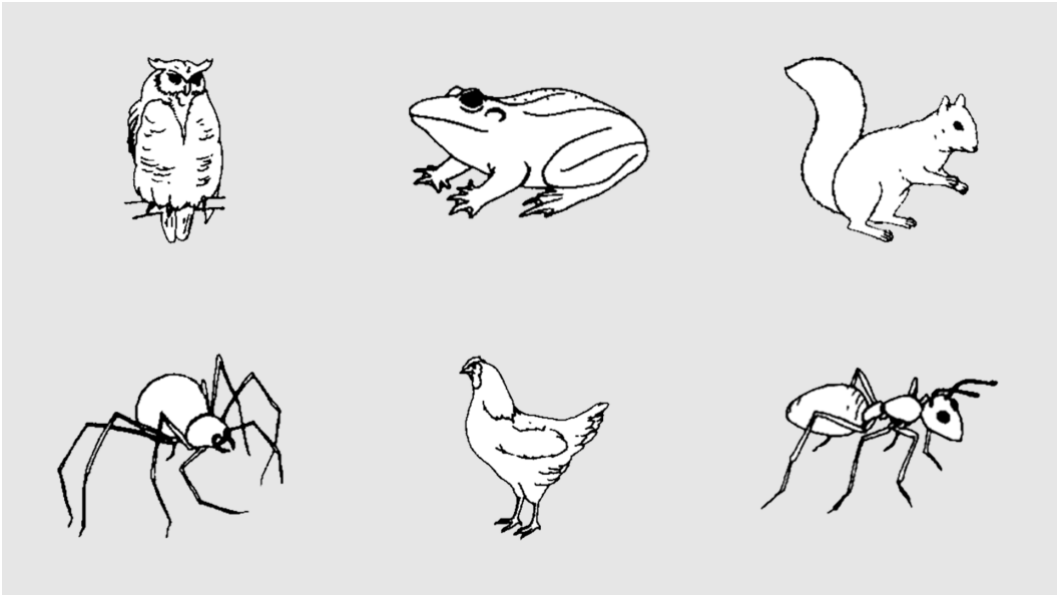


Table B2

Experimental Items for Figure B4

Subject Condition	Prepositional Predicative Condition
4.	*4.*
Q- *El BÚHO está arriba de la hormiga, ¿verdad?*	Q- *El búho está arriba de la HORMIGA, ¿verdad?*
[Q- The OWL is above the ant, right?]	[Q- The owl is above the ANT, right?]
A- *No, está la ARDILLA.*	A- *No, de la ARAÑA.*
[A- No, the SQUIRREL is].	[A- No, it is above the SPIDER].
5.	*5.*
Q- *La RANA está arriba de la araña, ¿verdad?*	Q- *La rana está arriba de la ARAÑA, ¿verdad?*
[Q- The FROG is above the spider, right?]	[Q- The frog is above the SPIDER, right?]
A- *No, está el BÚHO.*	A- *No, de la GALLINA.*
[A- No, the OWL is].	[A- No, it is above the HEN].
6.	*6.*
Q- *La ARDILLA está arriba de la gallina, ¿verdad?*	Q- *La ardilla está arriba de la GALLINA, ¿verdad?*
[Q- The SQUIRREL is above the hen, right?]	[Q- The squirrel is above the HEN, right?]
A- *No, está la RANA.*	A- *No, de la HORMIGA.*
[A- No, the FROG is].	[A- No, it is above the ANT].

Note. “Q” refers to the question asked by the experimenter and “A” to the expected response given by pre-schoolers.

Figure B5

Experimental Images for Table B3

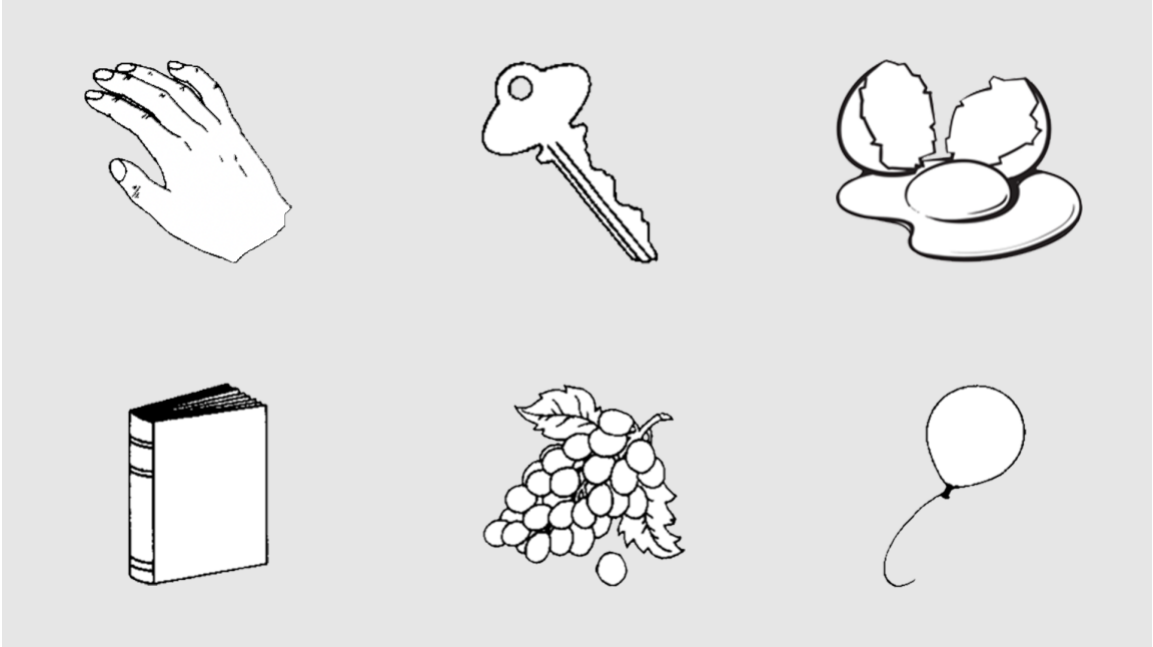


Table B3

Experimental Items for Figure B5

Subject Condition	Prepositional Predicative Condition
7.	7.
Q - *La MANO está arriba del globo, ¿verdad?*	Q- *La mano está arriba del GLOBO, ¿verdad?*
[Q- The HAND is above the balloon, right?]	[Q- The hand is above the BALLOON, right?]
A- No, está el HUEVO.	A- *No, del LIBRO.*
[A- No, the EGG is].	[A- No, it is above the BOOK].
8.	8.
Q- *La LLAVE está arriba del libro, ¿verdad?*	Q- *La llave está arriba del LIBRO, ¿verdad?*
[Q- The KEY is above the book, right?]	[Q- The key is above the BOOK, right?]
A- No, está la MANO.	A- *No, de la UVA.*
[A- No, the HAND is].	[A- No, it is above the GRAPE].
9.	9.
Q- *El HUEVO está arriba de la uva, ¿verdad?*	Q- *El huevo está arriba de la UVA, ¿verdad?*
[Q- The EGG is above the grape, right?]	[Q- The egg is above the GRAPE, right?]
A- No, está la LLAVE.	A- *No, del GLOBO.*
[A- No, the KEY is].	[A- No, it is above the BALLOON].

Note. “Q” refers to the question asked by the experimenter and “A” to the expected response given by pre-schoolers.

Figure B6

Experimental Images for Table B4

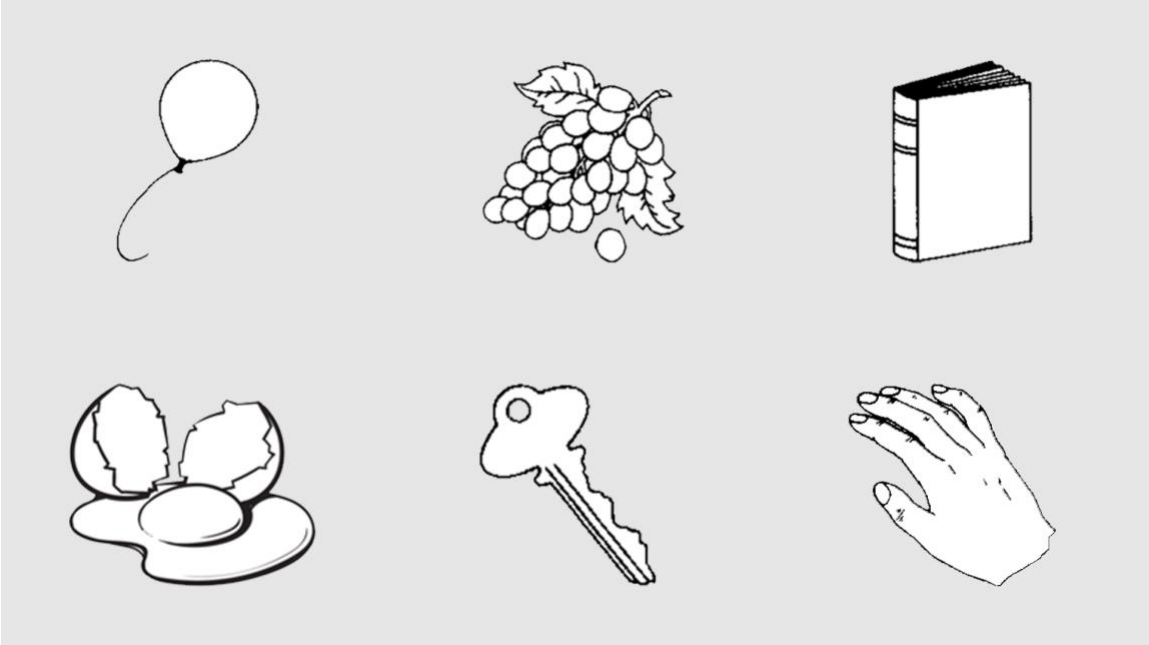


Table B4

Experimental Items for Figure B6

Subject Condition	Prepositional Predicative Condition
10.	10.
Q-El GLOBO está arriba de la mano, ¿verdad?	*Q-El globo está arriba de la MANO, ¿verdad?*
[Q- The BALLOON is above the hand, right?]	[Q- The balloon is above the HAND, right?]
A- No, está el LIBRO.	A- *No, del HUEVO.*
[A- No, the BOOK is].	[A- No, it is above the EGG].
11.	11.
Q-La UVA está arriba del huevo, ¿verdad?	*Q-La uva está arriba del HUEVO, ¿verdad?*
[Q- The GRAPE is above the egg, right?]	[Q- The grape is above the EGG, right?]
A- No, está el GLOBO.	A- *No, de la LLAVE.*
[A- No, the BALLOON is].	[A- No, it is above the KEY].
12.	12.
Q-El LIBRO está arriba de la llave, ¿verdad?	*Q-El libro está arriba de la LLAVE, ¿verdad?*
[Q- The BOOK is above the key, right?]	[Q- The book is above the KEY, right?]
A- No, está la UVA.	A- *No, de la MANO.*
[A- No, the GRAPE is].	[A- No, it is above the HAND].

Note. "Q" refers to the question asked by the experimenter and "A" to the expected response given by pre-schoolers.

Figure B7

Control Images for Table B5

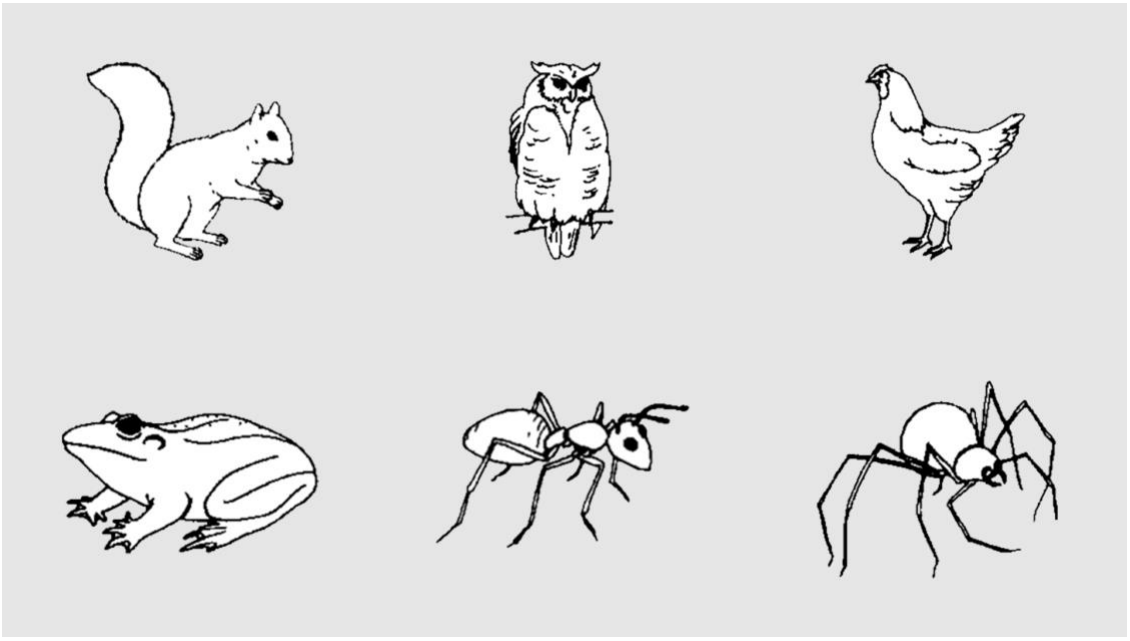


Table B5

Control Items for Figure B7

Subject Condition	Prepositional Predicative Condition
13.	131.
La ARDILLA está arriba de la rana,	*La ardilla está arriba de la RANA, ¿verdad?*
¿verdad?	[The BALLOON is above the hand, right?]
[The SQUIRREL is above the frog, right?]	
14.	14.
El BÚHO está arriba de la hormiga,	*El búho está arriba de la HORMIGA,*
¿verdad?	*¿verdad?*
[The OWL is above the ant, right?]	[The owl is above the ANT, right?]
15.	15.
La GALLINA está arriba de la araña,	*La gallina está arriba de la ARAÑA,*
¿verdad?	*¿verdad?*
[The HEN is above the spider, right?]	[The hen is above the SPIDER, right?]

Note. In these control items, all questions correctly describe the images.

Figure B8

Control Images for Table B6

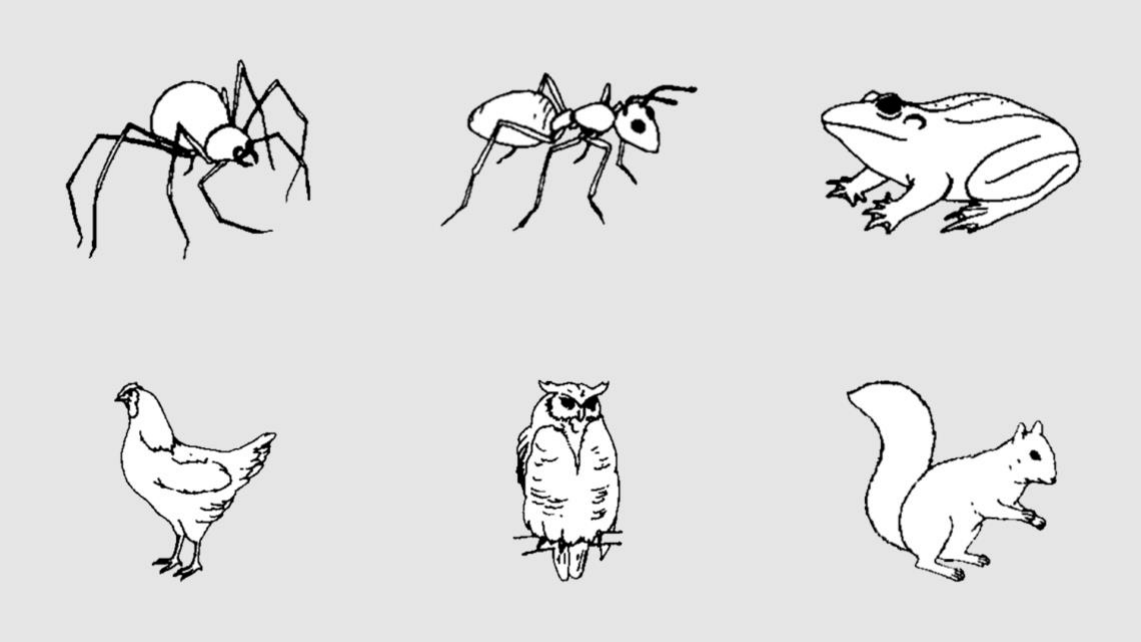


Table B6

Control Items for Figure B8

Subject Condition	Prepositional Predicative Condition
16.	16.
La ARAÑA está arriba de la gallina,	*La araña está arriba de la GALLINA,*
¿verdad?	*¿verdad?*
[The SPIDER is above the hen, right?]	[The spider is above the HEN, right?]
17.	17.
La HORMIGA está arriba del búho,	*La hormiga está arriba del BÚHO,*
¿verdad?	*¿verdad?*
[The ANT is above the owl, right?]	[The ant is above the OWL, right?]
18.	18.
La RANA está arriba de la ardilla, ¿verdad?	*La rana está arriba de la ARDILLA,*
[The FROG is above the squirrel, right?]	*¿verdad?*
	[The frog is above the SQUIRREL, right?]

Note. In these control items, all questions correctly describe the images.

Figure B9

Control Images for Table B7

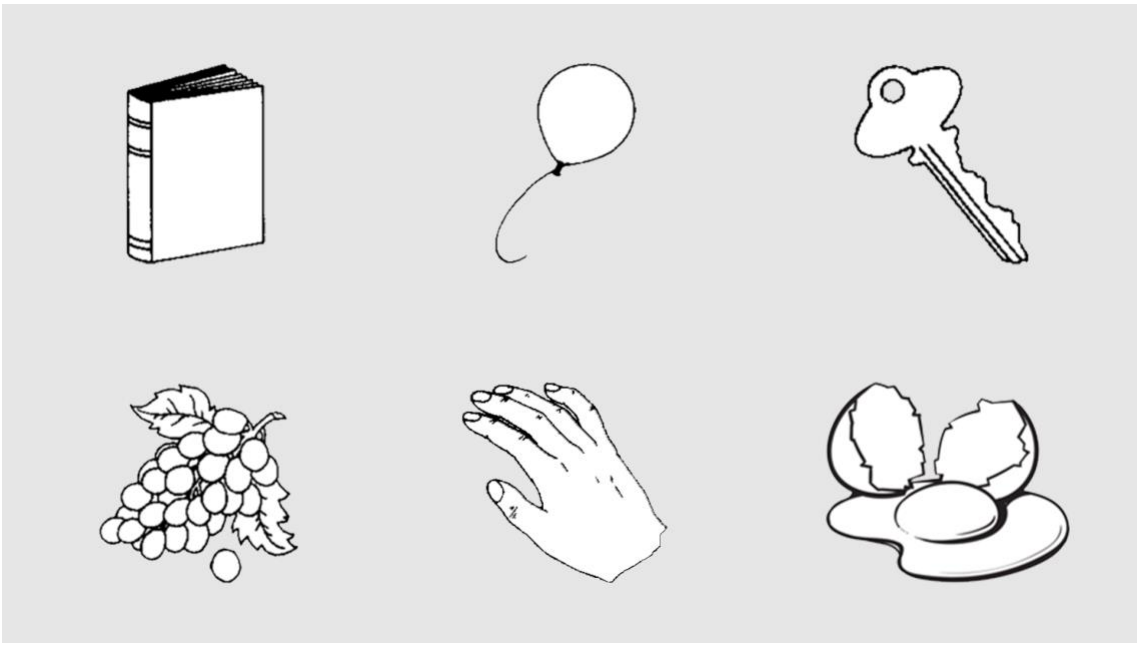


Table B7

Control Items for Figure B9

Subject Condition	Prepositional Predicative Condition
19.	19.
El LIBRO está arriba de la uva, ¿verdad?	*El libro está arriba de la UVA, ¿verdad?*
[The BOOK is above the grape, right?]	[The book is above the GRAPE, right?]
20.	20.
El GLOBO está arriba de la mano, ¿verdad?	*El globo está arriba de la MANO, ¿verdad?*
[The BALLOON is above the hand, right?]	[The balloon is above the HAND, right?]
21.	21.
La LLAVE está arriba del huevo, ¿verdad?	*La llave está arriba del HUEVO, ¿verdad?*
[The KEY is above the egg, right?]	[The key is above the EGG, right?]

Note. In these control items, all questions correctly describe the images.

Figure B10

Control Images for Table B8

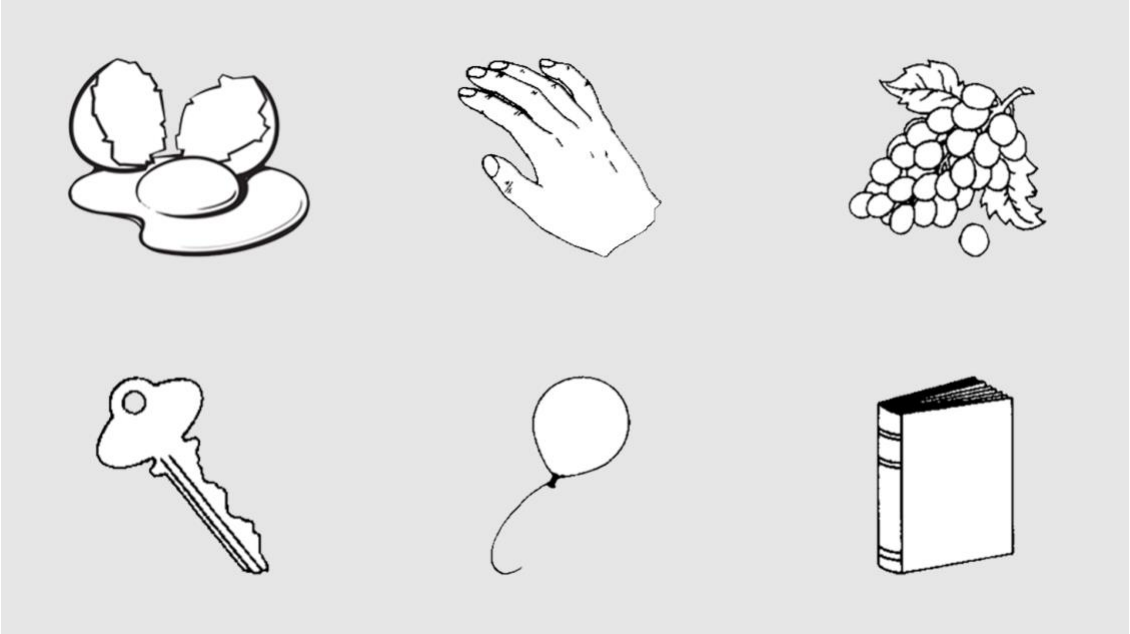


Table B8

Control Items for Figure B10

Subject Condition	Prepositional Predicative Condition
22.	22.
El HUEVO está arriba de la llave, ¿verdad?	*El huevo está arriba de la LLAVE, ¿verdad?*
[The EGG is above the key, right?]	[The egg is above the KEY, right?]
23.	23.
La MANO está arriba del globo, ¿verdad?	*La mano está arriba del GLOBO, ¿verdad?*
[The HAND is above the balloon, right?]	[The hand is above the BALLOON, right?]
24.	24.
La UVA está arriba del libro, ¿verdad?	*La uva está arriba del LIBRO, ¿verdad?*
[The GRAPE is above the book, right?]	[The grape is above the BOOK, right?]

Note. In these control items, all questions correctly describe the images.

Figure B11

Filler Images for Table B9

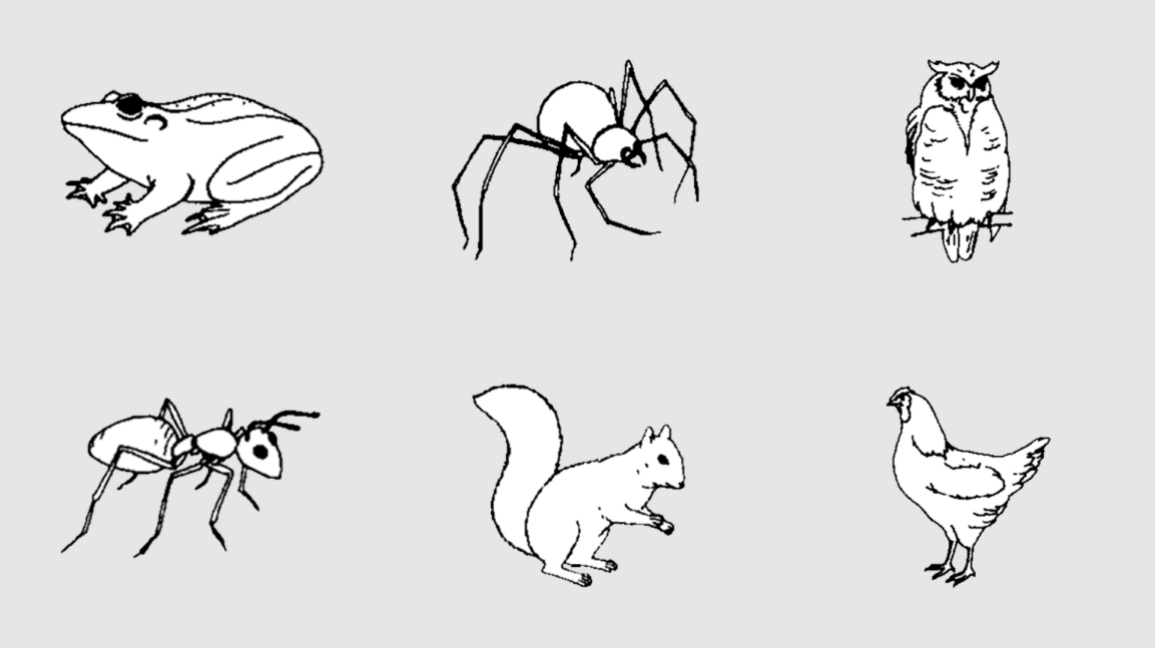


Table B9

Filler Items for Figure B11

Right Answer	Wrong Answer
25.	25.
La hormiga está abajo de la rana, ¿verdad?	*La hormiga está abajo del búho, ¿verdad?*
[The ant is above the frog, right?]	[The ant is above the owl, right?]
26.	26.
La ardilla está abajo de la araña, ¿verdad?	*La ardilla está abajo de la rana, ¿verdad?*
[The squirrel is above the spider, right?]	[The squirrel is above the frog, right?]
27.	27.
La gallina está abajo del búho, ¿verdad?	*La gallina está abajo de la araña, ¿verdad?*
[The hen is above the owl, right?]	[The hen is above the spider, right?]

Note. In these filler items, half of the questions correctly describe the images and half of them do not. These utterances have a neutral intonation.

Figure B12

Filler Images for Table B10

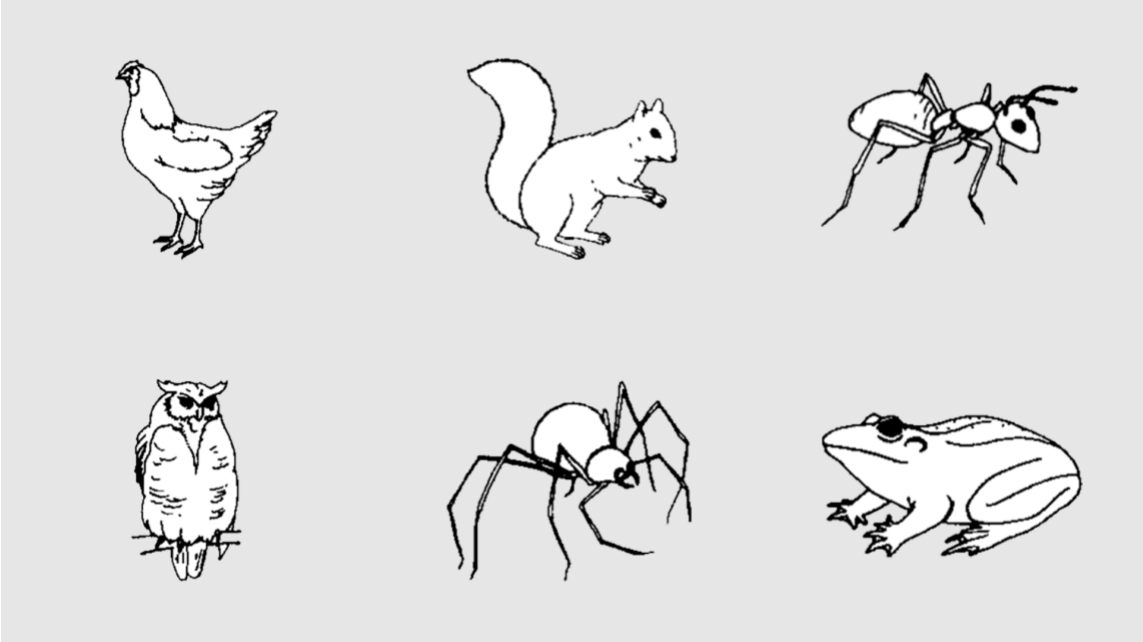


Table B10

Filler Items for Figure B12

Right Answer	Wrong Answer
28.	28.
El búho está abajo de la gallina, ¿verdad?	*El búho está debajo de la hormiga, ¿verdad?*
[The owl is above the hen, right?]	[The owl is above the ant, right?]
29.	29.
La araña está abajo de la ardilla, ¿verdad?	*La araña está abajo de la gallina, ¿verdad?*
[The spider is above the squirrel, right?]	[The spider is above the hen, right?]
30.	30.
La rana está abajo de la hormiga, ¿verdad?	*La rana está abajo de la ardilla, ¿verdad?*
[The frog is above the ant, right?]	[The frog is above the squirrel, right?]

Note. In these filler items, half of the questions correctly describe the images and half of them do not. These utterances have a neutral intonation.

Figure B13

Filler Images for Table B11

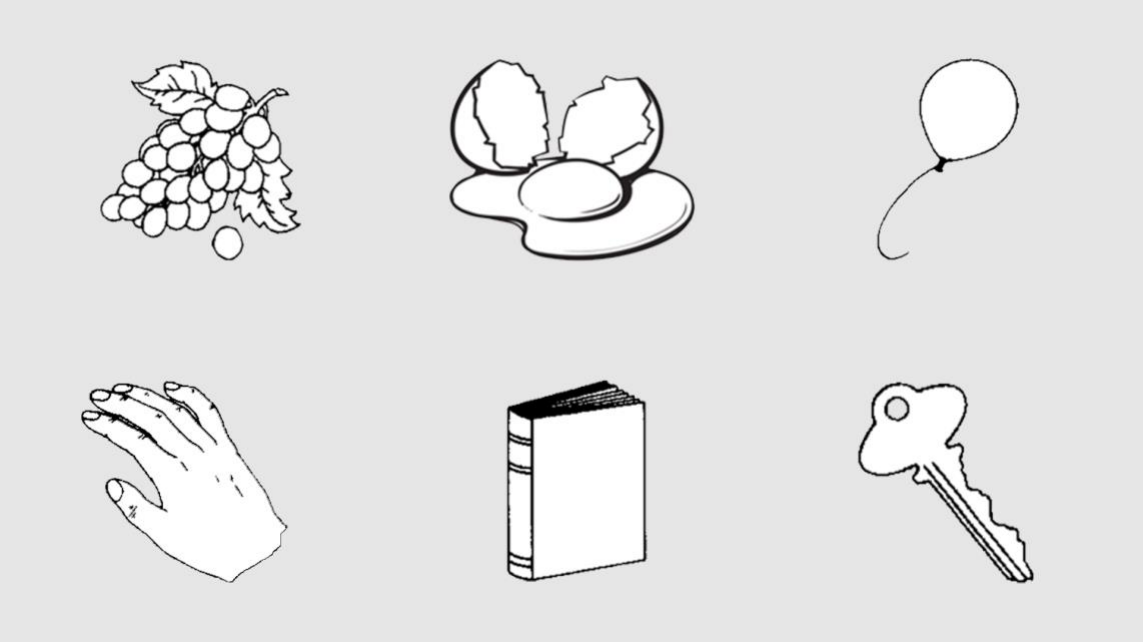


Table B11

Filler Items for Figure B13

Right Answer	Wrong Answer
31.	31.
La mano está abajo de la uva, ¿verdad?	*La mano está abajo del globo, ¿verdad?*
[The hand is above the grape, right?]	[The hand is above the balloon, right?]
32.	32.
El libro está abajo del huevo, ¿verdad?	*El libro está abajo de la uva, ¿verdad?*
[The book is above the egg, right?]	[The book is above the grape, right?]
33.	33.
La llave está abajo del globo, ¿verdad?	*La llave está abajo del huevo, ¿verdad?*
[The key is above the balloon, right?]	[The key is above the egg, right?]

Note. In these filler items, half of the questions correctly describe the images and half of them do not. These utterances have a neutral intonation.

Figure B14

Filler Images for Table B12

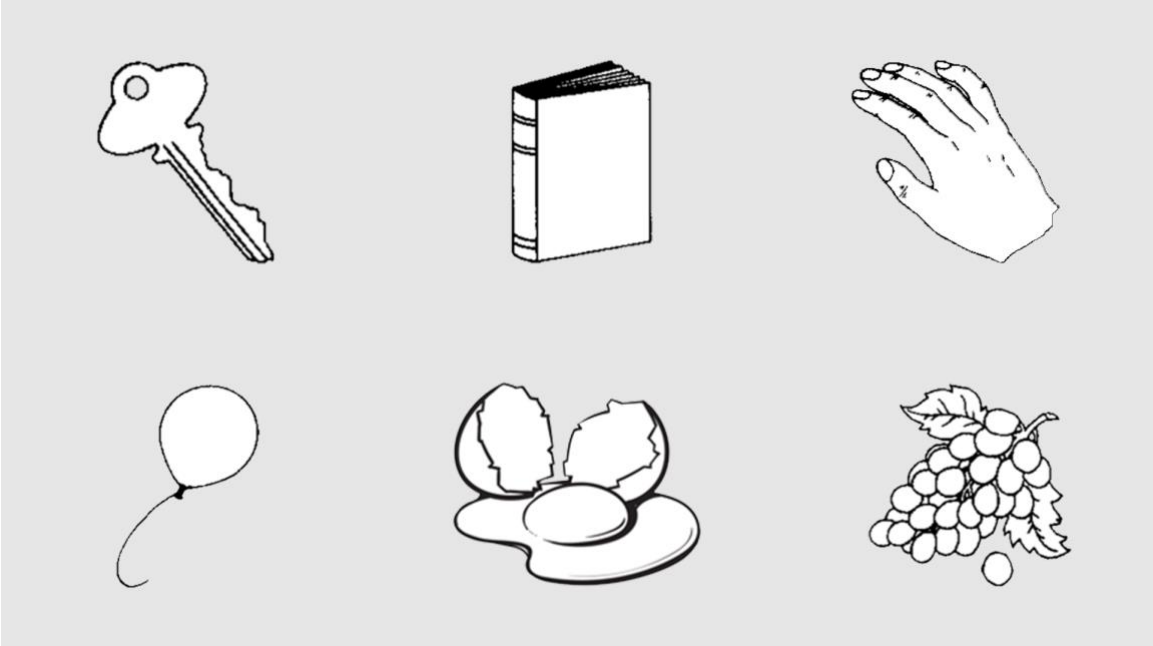


Table B12

Filler Items for Figure B14

Right Answer	Wrong Answer
34.	34.
El globo está abajo de la llave, ¿verdad?	*El globo está abajo de la mano, ¿verdad?*
[The balloon is above the key, right?]	[The balloon is above the hand, right?]
35.	35.
El huevo está abajo del libro, ¿verdad?	*El huevo está abajo de la llave, ¿verdad?*
[The egg is above the book, right?]	[The egg is above the key, right?]
36.	36.
La uva está abajo del mano, ¿verdad?	*La uva está abajo del libro, ¿verdad?*
[The grape is above the hand, right?]	[The grape is above the book, right?]

Note. In these filler items, half of the questions correctly describe the images and half of them do not. These utterances have a neutral intonation.

Appendix C

Creation of Auditory Stimuli

Auditory stimuli were recorded by a female native Mexican Spanish speaker from Mexico City. To record the filler items, she was instructed to read a list of sentences without any context, to eliminate any bias where narrow prosodic focus may be used. For this reason, these utterances were recorded first, before showing any pictures. An example of this list presented is shown in Figure C1.

Figure C1

Example of the List of Filler Items as Presented to the Native Speaker

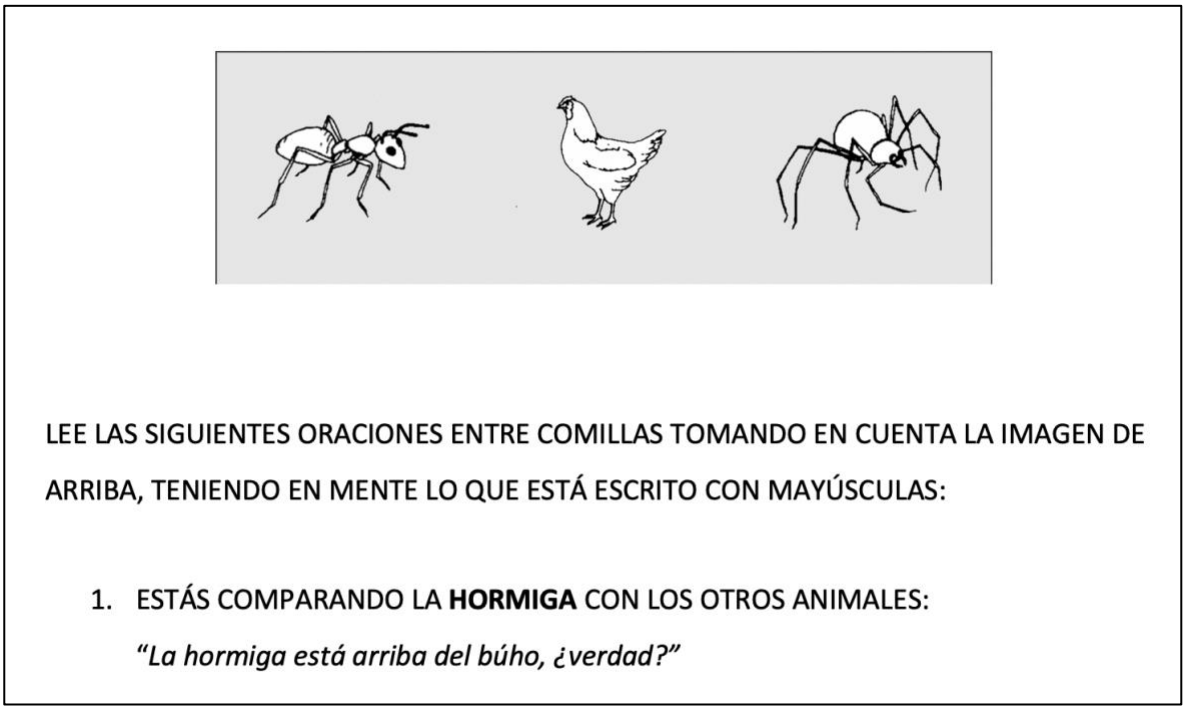
1. *"La hormiga está abajo de la rana, ¿verdad?"*
2. *"La ardilla está abajo de la araña, ¿verdad?"*
3. *"La gallina está abajo del búho, ¿verdad?"*
4. *"La hormiga está abajo del búho, ¿verdad?"*
5. *"La ardilla está abajo de la rana, ¿verdad?"*
6. *"La gallina está abajo de la araña, ¿verdad?"*
7. *"El búho está abajo de la gallina, ¿verdad?"*
8. *"La araña está abajo de la ardilla, ¿verdad?"*

Then, the Subject Condition (SC) and Prepositional Predicative Condition (PPC) were recorded one after the other. For this, images with only the compared subjects or prepositional predicatives were presented along with the sentences, without the rest of the items that were

below or above them. This way, the speaker did not know if the statements said about the images (what element is above or below another) were correct or not, to avoid any bias this knowledge may cause in the production of the sentences. An example of how these elements were presented is shown in Figure C2.

Figure C2

Example of the List of Experimental and Control Items as Presented to the Native Speaker



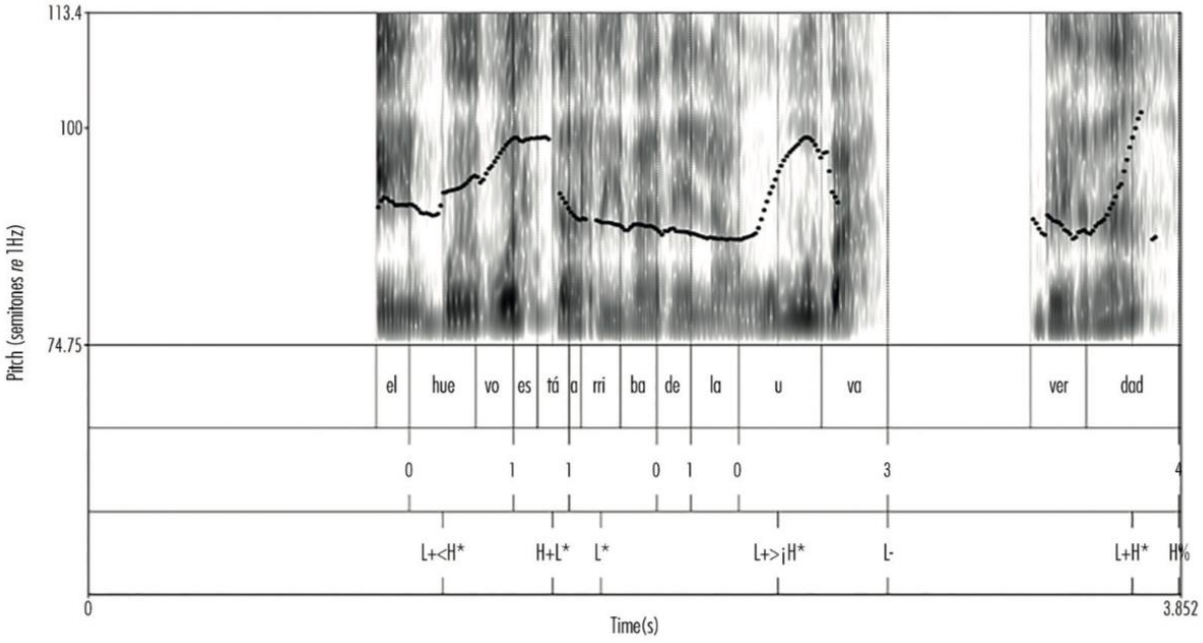
Note. The instructions are translated as follows: "Read the following sentences within quotation marks taking into account the image above, having in mind what is written in capital letters". The instructions before the sentence itself state, in this example: "YOU ARE COMPARING THE **ANT** WITH THE REST OF THE ANIMALS".

The utterances produced were controlled in Praat (Boersma & Weenink, 2006) to ensure they were similar in each condition. To guarantee the natural quality of the speaker's production, these stimuli were not manipulated; however, the utterances that kept a consistent prosodic focus marking in each condition were selected.

In the case of the PPC, it was observed that the speaker produced an elevation of the intonational curve in the accented syllable of the prepositional predicatives of 8.3 semitones (st) on average. The tone pattern observed was L+>H* in most cases, with the intonational curve reaching its peak inside this syllable in all cases. In the subject of the sentences, there was an elevation of the intonational curve in its tonic syllable of 2.5 st on average. The peak of the F0 was in the posttonic syllable, with a tone pattern of L+<H* in all cases. Additionally, the tonic syllables of the prepositional predicatives had a greater duration ($\bar{x}$=295 milliseconds, or ms) than the tonic syllables of the subjects ($\bar{x}$=168 ms). This could be interpreted as a greater prosodic prominence on the prepositional predicatives than on the subjects in this condition (Hualde, 2014). An example of stimuli of the PPC is shown in Figure C3:

Figure C3

Example of Audio Stimuli of the Prepositional Predicative Condition



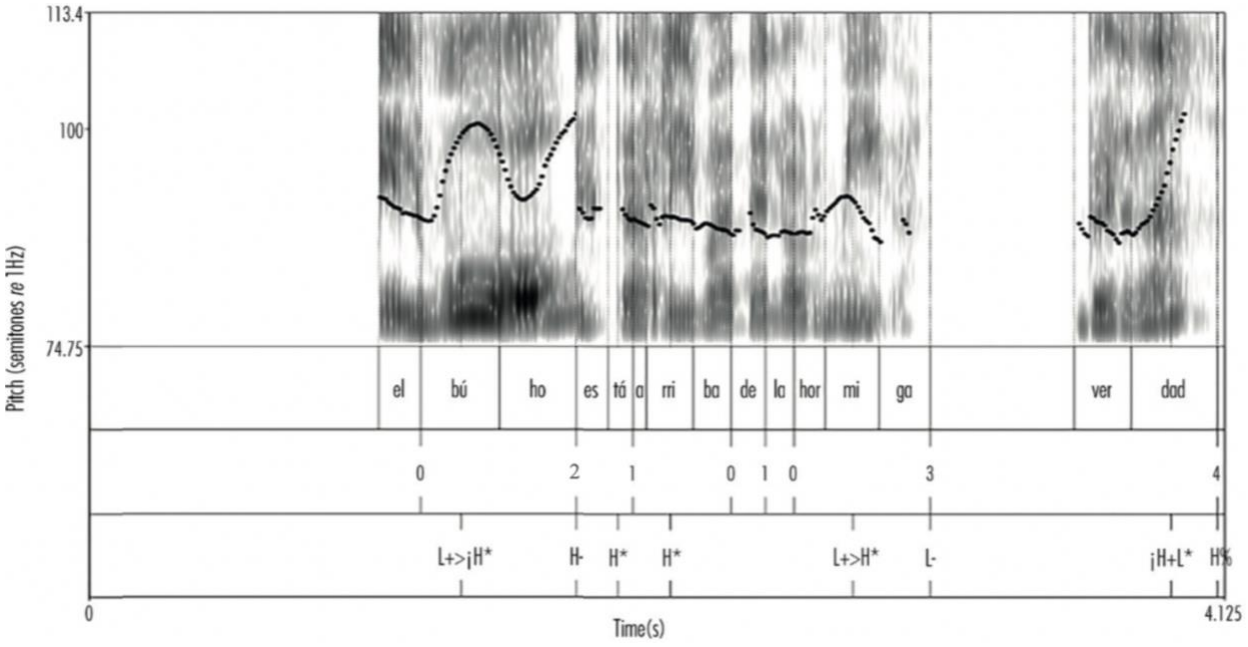
In the case of the SC, there was an elevation of the intonational curve in the tonic syllable of the subjects of 6.1 st on average. The peak of the F0 was in the tonic syllable, with a tone pattern of L+>¡H* in all cases. In addition, the prepositional predicative of the sentences had a H* or L+>H* tone pattern, with an elevation of the intonational curve of 1.8 st on average. Additionally, there was a greater duration of the tonic syllables of the subjects ($\bar{x}$=246 ms) than the tonic syllables of the prepositional predicatives ($\bar{x}$=188 ms). This could be interpreted as a greater prosodic prominence on the subjects than on the prepositional predicatives in this condition (Hualde, 2014).

Furthermore, a pause was observed between the subject and the rest of the sentence in this condition, with a H- boundary tone; this pause was eliminated because it was an additional cue

of prosodic focus in this condition that was not present in the PPC. An example of the stimuli of the SC is shown in Figure C4:

Figure C4

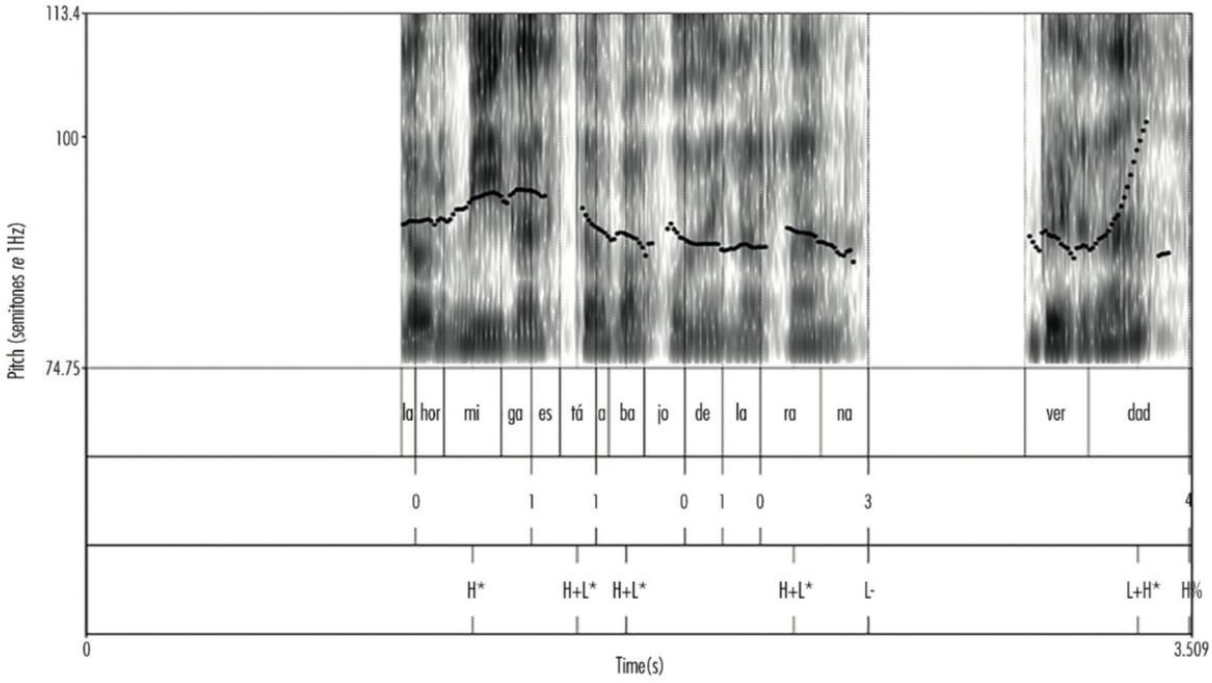
Example of Audio Stimuli of the Subject Condition



Finally, filler items show elevations of the intonational curve of 1.5 st in the subject and 1.6 st in the prepositional predicative on average. In addition, when it comes to the subject, the peak of the F0 is in the posttonic syllable, and the intonational curve falls progressively, which are characteristics of the declarative sentences with wide focus (when focus is not on a particular word of an utterance) (Hualde, 2014). The tonic syllables had a short duration, 180 ms on average in the case of the subject and 170 ms on average in the case of the prepositional predicative. An example of filler items is shown in Figure C5:

Figure C5

Example of Audio Stimuli of the Filler Items



Appendix D

Auditory Stimuli Validation

A total of twenty adults between the ages of 24 and 40 years ($\bar{x}$=31;6, *SD*=8.8) from the central and central-western Mexico took part in the experiment. Participants were selected to each condition randomly (SC, *n*=11; PPC, *n*=9). They accessed the experiment through a link to the platform "Gorilla" (https://gorilla.sc/) that was shared on social media.

The experiment consisted of the 36 experimental trials used in the Experiments 1 and 2, with 12 experimental items, 12 control items and 12 filler items.

Procedure of Auditory Stimuli Validation

Before starting the task, a consent form was presented, in which they were instructed to check a box to agree to take part of the experiment. Then, a questionnaire was given, asking information about their date of birth, the state and country they were from and if they had any hearing impairments. After this, the task began, preceded by a text and recording on screen that instructed them to enter full screen mode. Later, the instructions were presented by text and recordings as well. These instructions were as follows:

"*A continuación verás una serie de imágenes seguidas de una afirmación sobre ellas. Deberás indicar si la afirmación es CORRECTA o INCORRECTA. Si es INCORRECTA, indica la respuesta correcta en voz alta*" ["You will now see a series of images followed by a statement about them. You must indicate if the statement is CORRECT or INCORRECT. If it is INCORRECT, state the correct answer out loud."]

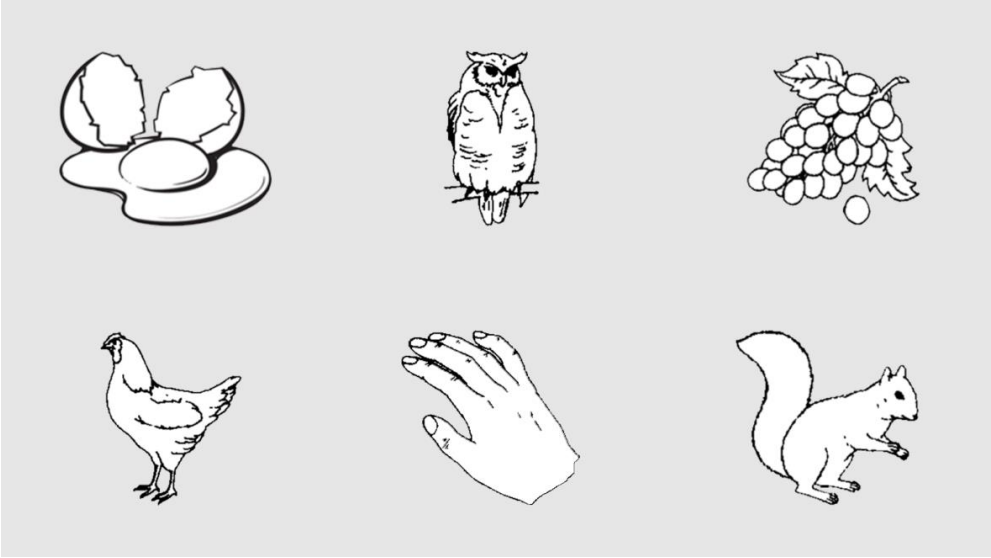
Then, a test recording was performed with the following instructions: "*Al dar la respuesta debes hablar en voz alta. Lo que respondas será grabado. Por favor, autoriza el uso del micrófono en tu dispositivo. Para asegurarnos de que el micrófono funcione bien, realiza una*

prueba de grabación. Haz click en 'Grabar' para empezar la grabación, y en 'Detener grabación' para finalizarla. Puedes reproducir tu grabación para asegurarte de que se escuche correctamente" ["When you give the answer you should speak out loud. Whatever you answer will be recorded. Please authorize the use of the microphone on your device. To make sure the microphone is working well, please do a test recording. Click on 'Record' to start recording, and on 'Stop recording' to end the recording. You can play your recording to make sure it is recorded correctly."]

After this, an example of the task was given. For this example, the following image (Figure D1), which contained animate and inanimate objects at the same time and was not part of the experimental images, was presented:

Figure D1

Example Image of the Auditory Stimuli Validation



To avoid biased results, the example of what to respond if an incorrect statement was given was not provided. The instructions stated:

"*Por ejemplo, verás una imagen como la siguiente, y escucharás algo como: 'El huevo está a un lado de la mano, ¿verdad?'. Si la respuesta es correcta, menciona "Sí" o "Correcto". Si la respuesta es incorrecta, responde con la respuesta correcta. ¿En este caso, cuál es? Di en voz alta la respuesta correcta a continuación*" ["For example, you will see an image like the following, and you will hear something like: 'The egg is on one side of the hand, right?'. If the answer is correct say out loud "Yes" or "Correct". If the answer is incorrect, provide the correct response. In this case, what is it? Say out loud the correct answer below."]

It was expected that the participants would notice that the statement about the image was incorrect and hence gave an answer that described the image the best.

Finally, the task began with the following remainder: "*RECUERDA: Si la afirmación es INCORRECTA, menciona la respuesta correcta en voz alta. Responde con voz clara y fuerte; después de 7 segundos se pasará a la siguiente pregunta.¡EMPECEMOS!"* ["REMEMBER: If the statement is INCORRECT, state the correct answer out loud. Respond in a clear and loud voice; after 7 seconds, the task will proceed with the next question. LET'S BEGIN!"].

As stated by the instructions, after the presentation of each stimulus, 7 seconds were given to say the answer. This time was established after test runs with members of the laboratory (n= 5), where 5 seconds proved too short a time to answer and 10 seconds, too long. This measure was applied with adults only and was not used for a children population.

The answers given were recorded for further analysis. Only the recordings from the experimental and control stimuli (a total of 24 trials) were analysed, as these were the stimuli of

interest for the experiment, where wrong statements were heard and the participants were expected to correct them, showing their understanding of prosodic focus.

Results of Auditory Stimuli Validation

A total of 684 audio recordings were collected (SC, *n*=288; PPC, *n*=396). Data cleansing was performed, in which 22 audio recordings were inaudible in the PPC. In addition, 6 recordings had incomplete responses (SC, *n*=2; PPC, *n*=4). This ended up with a total of 428 recordings for the analysis (SC, *n*=190; PPC, *n*=238).

The first analysis consisted of assessing the correct responses, with a total of 173 recordings (CS, *n*=69; PPC, *n*=104) answering to the control and experimental stimuli. In this case, a response was classified as "correct" only if the participants answered with a correction to the experimenter: a "no" or other negative answer to the experimental stimuli; and "yes", "correct" or other affirmative answer to the control stimuli. The results showed that the prepositional predicative was corrected in the PPC 100% of the times; while, in the SC, the subject was corrected only 14.49% of the times. Therefore, the SC participants mostly corrected the prepositional predicative of the sentence, giving identical answers to the PPC participants. This could be explained by prosody being the only cue to mark focus in the experimental stimuli, while syntactic movement is used in Spanish to mark prosodic focus when it is contrastive or exhaustive (when it makes comparisons between elements), as well as the canonical position of focus being at the end of the phrases in this language (Szendröi et al., 2018).

However, listening to the recordings of the validation, even if most of the adults corrected the prepositional predicative in the SC, giving identical answers to the ones in the PPC, they seemed to emphasize the subject of their utterances. A greater emphasis on the prepositional predicative also seemed present in the answers to the PPC. According to Wilson and Wharton (2006), a

greater articulatory effort in an element of an utterance is always taken into consideration by a hearer as a relevant cue to interpret the sentence, as it deviates from intonation without this unexpected effort in the rest of the elements of the phrase. With emphasis to different elements depending on the condition, participants seemed to perceive the articulatory effort used to emphasize the subject in the SC and the prepositional predicative in the PPC.

To prove this perception of prosodic cues, an additional analysis was made with the recordings of participants that gave identical responses in the SC and PPC, under the hypothesis that they correspondingly emphasised the subject or the prepositional predicative of the phrases (for example, “No, the SQUIRREL is above the frog” in the SC and “No, the squirrel is above the FROG” in the PPC). A total of 82 recordings from the PPC and 34 from the SC were analysed. Only responses where the whole phrase was corrected were analysed to compare emphasis in the subjects and prepositional predicatives of the utterances (answers like “No, the squirrel is above the frog”, for example, instead of just “No, above the frog”).

A comparison on the tone pattern was initially made between the subject and the prepositional predicative of the utterances in both conditions. This acoustic analysis was done using Praat (Boersma & Weenink, 2006), the Sp_ToBI (Tones and Breaks Indices) transcription system (Beckman et al., 2002) and the prosodic hierarchy indices by Nespor and Vogel (1994) described in the section Results of Experiment 1. The results showed a different tone pattern between conditions regarding the subject of the sentence, as shown in Table D1 below:

Table D1

Tones Observed in the Responses Given by Pre-schoolers on the Pilot Test

Tones on the Subject		**Tones on the Prepositional Predicative**	
Subject Condition	Prepositional Predicative Condition	Subject Condition	Prepositional Predicative Condition
L+<H* (44.11%)	L+<H* (67%)	L* (44.11%)	L+>H* (39%)
L+>H* (23.5%)	H* (28%)	L+>H* (24.5%)	L* (31.7%)
L+H* (17.6%)	L+H* (2.4%)	H+L* (11.7%)	H+L* (20.7%)
H* (11.8%)	H+L* (1.2%)	L+H* (8.8%)	H* (4.9%)
H+L* (2.9%)	H+>L* (1.2%)	H* (8.8%)	L+H* (3.6%)

In the subject of the utterance, the most used tone pattern was L+<H* on both conditions, which is usually used in Spanish to mark the topic of the sentence (Hualde, 2013). However, the second most common tone pattern in the SC was L+>H*, used in Spanish to mark prosodic focus (Hualde, 2013). This was also the most common tone pattern on the prepositional predicative of the sentence on the PPC; different to the L* in the SC which does not mark emphasis. Nevertheless, these two different tone patterns were common in both conditions.

Additionally , regarding intermediate junction tones, there was a pause between the subject and the rest of the statement with tone H- only in the SC in 38.23% of the cases; although a pause was perceived in both conditions after the subject as there was an abrupt fall in the melodic curve, most after a L+<H* bitone, in 26% of the cases. Even so, the pause perceived in the SC was more prominent.

Furthermore, the pitch height, the elevation in semitones of the melodic curve and the duration of the accented syllable of the subject and prepositional predicative of the sentence in both conditions were also analysed. A Mann-Whitney *U* test was performed because of the small sample size.

The results showed a significative difference between conditions only in the subject of the sentence, regarding pitch elevation (U= 919, p= 0.004), with a greater elevation in the SC ($\bar{x}$= 2.98 st, SD= 1.39) than in the PPC ($\bar{x}$= 2.25 st, SD=1.25). Additionally, there was a greater pitch height on the accented syllable in the SC ($\bar{x}$= 95.42 st, DE= 2.07) in comparison with the PPC ($\bar{x}$= 91.58 st, DE= 5.11), with a significative difference between conditions (U= 706, p< 0.001). No significant difference was found in the duration of the subject's stressed syllable (U= 1093, p= 0.068); or in the pitch (U= 1085, p= 0.061), elevation of the melodic curve (U= 1181, p = 0.196) or duration of the stressed syllable (U= 1147, p= 0.134) of the prepositional predicatives of the sentence.

Taking the results of the acoustic analysis all together, it seems that Spanish adult speakers distinguish prosodic differences between conditions in the experiment. They make a greater distinction in the prepositional predicative of the sentence by emphasizing it in the PPC with a tone L+>H*. In addition, they also distinguish the subject with a greater elevation of the melodic curve and pitch in the subject in the SC. In conclusion, they distinguish both conditions with different acoustic features.

However, unless there is a clear context that stablishes that the speaker (the experimenter) and the hearer (the participants) of a conversation (the experiment) share the same expectations, it is difficult to conclude that there was a clear comprehension of prosodic focus in the experimental task. In this validation, utterances were presented in isolation, with the sole instruction to correct wrong statements. Wilson and Wharton (2006) also noted that the expectations of the hearer must always be considered, creating a context rich enough so that acoustic prominence implicates a strong conclusion. Isolated phrases with acoustic prominence in one word can lead to different conclusions, and hence the conclusion desired will be weakly implicated; for this

reason, it is essential to create a clear context that shows the intentions of the speaker. Prosodic cues alone are not enough if speaker and hearer do not share the same context or expectations. In this sense, as stated by Villalobos-Pedroza (2021), prosodic cues in focus can also be prosodic cues used in other elements of speech with different pragmatic meanings; the context of the conversation is what gives prosodic cues in focus its different and distinctive communicative functions.

In consequence, considering the importance of context as a common base between speaker and hearer in a conversation, it was fundamental that the experiment established the comparison between components, which is the main goal of the experiment, as well as the function of the focus used in the tasks. Subsequently, the Experiments 1 and 2 were presented as a game in which the participants must help a character to remember the correct position of the images on screen, comparing two of them at a given time.

Appendix E

Pilot Test of Experiment 1

A pilot testing was performed with 8 Spanish-speaking pre-schoolers of 4 to 5 years of age ($\bar{x}$ = 5;2, SD= 0.36) from central-western Mexico, a dialectal zone proposed by Martín-Butragueño (2014): 5 from Aguascalientes, 1 from Mexico City, 1 from Jalisco and 1 from Morelos. The participants were randomized in both conditions (SC, *n*= 4; PPC, *n*= 4). They did not have any language, cognitive or auditory impairments as reported by their parents.

The procedure was as described in the section Procedure of Experiment 1. A total of 288 audio files (144 per condition) were obtained. From this total, only the recordings that belonged to the control stimuli (where the participants were expected to answer with "yes" or "correct" to the correct statements about the images) and to the experimental stimuli (where the participants had to correct the wrong statement said by the experimenter) were considered for the analysis. This left a total of 196 audio recordings (96 per condition).

The first analysis consisted of the percentage of correct responses. In this case, a response was classified as "correct" only if the participants answered with a correction to the experimenter, a "no" or other negative answer to the experimental stimuli; and "yes", "correct" or other affirmative answer to the control stimuli. In this sense, there was a total of 85.45% of correct responses to the SC and a total of 93.75% correct answers to the PPC. In consequence, the percentage of correct answers was above chance in both cases, and hence the experiment was appropriate for pre-schoolers of 4 to 5 years of age.

The second analysis consisted of comparing the answers given to the experimental stimuli in both conditions. These answers consisted of 48 recordings per condition. It was expected that the pre-schoolers would correct the subject in the SC (for example, correct which animal is above

the squirrel in “The SPIDER is above the squirrel, right?”) and the prepositional predicative in the PPC (in this case, correct which animal is below the spider in “The spider is above the SQUIRREL, right?”).

The results showed that there was a total of 4 corrections to the subject in the SC and 44 to the prepositional predicative. In contrast, the PPC had 31 to the prepositional predicative and 17 to the subject. There was a significant 2x2 association between condition and corrected word, $x^2(1)= 10.301$, p= 0.001. This was a medium effect (Cramer’s V=0.328). Residual analysis showed that the subject had more corrections in the PPC than would be expected by chance (adjusted residual= 3.2; 35.5% in the PPC and 8.3% in the SC); whilst the prepositional predicative had more corrections in the SC than would be expected by chance (adjusted residual= 3.2; 91.7% in the SC and 64.6% in the PPC).

These results are the opposite pattern to what was expected: there was a greater number of corrections to the subject in the PPC and to the prepositional predicative in the SC. It is possible that the emphasized word is what stays in the pre-schoolers’ memory and hence becomes the topic of their corrections, focusing on correcting the other element that does not have prosodic prominence. However, the results of the experiment itself, with a greater number of participants, would help to have a clearer picture.

Even so, there were more corrections, in general, to the prepositional predicative of the sentence in both conditions. Because of this, a last analysis was performed with sentences that had the same syntactic structure in both conditions. An acoustic analysis was performed to analyse if there was a difference acoustically between them, as it was perceived that, even if the answers were the same in both conditions, correcting the prepositional predicative, emphasis was made by the participants on the subject in the SC and on the prepositional predicative in the PPC.

These recordings had answers where the pre-schoolers corrected the experimenter uttering a sentence with subject, verb, and object configuration (for example, “No, the squirrel is above the owl”). For this, 24 sentences were selected in the PPC and 12 in the CS (from the total of 48 per condition) with these characteristics.

A comparison on the tone pattern was made first between the subject and the prepositional predicative of the utterances in both conditions. The acoustic analysis was made using Praat (Boersma & Weenink, 2006), using the Sp_ToBI (Tones and Breaks Indices) transcription system (Beckman et al., 2002) and the prosodic hierarchy indices by Nespor and Vogel (1994) described in Results of Experiment 1. The results showed a different tone pattern between conditions regarding the subject of the sentence, as shown Table E1 below:

Table E1

Tones Observed in the Responses Given by Pre-schoolers on the Pilot Test

Tones on the Subject		**Tones on the Prepositional Predicative**	
Subject Condition	Prepositional Predicative Condition	Subject Condition	Prepositional Predicative Condition
H+L* (50%)	L+<H* (87.5%)	L+>H* (41.66%)	L+H* (29.16%)
L+<H* (16.66%)	H* (12.5%)	L+H* (41.66%)	L+>H* (25%)
L+H* (8.33%)		L* (8.33%)	H* (16.66%)
L* (8.33%)		H* (8.33%)	L* (16.66%)
H* (8.33%)			H+L* (12.5%)
L+>H* (8.33%)			

The most used tone pattern was L+<H* on the PPC, which is usually used in Spanish to mark the topic of the sentence (Hualde, 2013); this was not the case of the SC, where the most common pattern was H+L* in the subject, something that was not reported before. The most

common tone patterns on the prepositional predicative of the sentence were the same in both conditions, L+>H* and L+H*, both usually used to mark focus in Spanish (Hualde, 2013).

Additionally, regarding intermediate junction tones, there was a pause between the subject and the rest of the statement with tone H- only in the SC in 41.66% of the cases; although a pause was perceived in both conditions after the subject as there was an abrupt drop in the melodic curve, without distinction in pitch accent, in 33.33% of the cases. Even so, the real pause in the SC was more prominent.

Moreover , the tone height, the elevation in semitones of the melodic curve and the duration of the accented syllable of the subject and prepositional predicative of the sentence in both conditions were also analysed. A Mann-Whitney *U* test was performed because of the small sample size. The results showed a significative difference between conditions regarding the tone elevation in the subject (U= 68, p= 0.011), with a greater elevation in the SC ($\bar{x}$= 4.74 st, SD= 2.88) than in the PPC ($\bar{x}$= 2.34 st, SD=1.07). Additionally, there was a greater duration on the accented syllable of the subject in the SC ($\bar{x}$= 269.83 ms, DE= 0.04) in comparison with the PPC ($\bar{x}$= 230.04 ms, DE= 0.07), with a significative difference between conditions (U= 76, p= 0.022).

No significant difference was found in the pitch of the subject (U= 132.50, p= 0.699), nor in the pitch (U= 100.50, p= 0.144) or duration (U= 140, p= 0.893) of the prepositional predicative of the sentence; only a marginally significant difference in the prepositional predicative in the semitone elevation of the melodic curve (U= 87, p= 0.058), with a mean of 3.8 st in the SC (SD= 2.34) against a mean of 2.35 st. in the PPC (SD= 1.52).

Taking the results of the acoustic analysis all together, it seems that Spanish-speaking pre-schoolers can distinguish the prosodic differences between conditions, especially in the subject of the sentences. The tone H+L* found in the subject of the SC could be explained as it being

characteristic of the end of the phrase, and reflective of the pause between the subject and the rest of the statement, which would also indicate a type of emphasis commonly used for contrastive focus (Martín Butragueño, 2005). This emphasis in the subject, therefore, could have been more salient as it is a position where focus is not expected in Spanish. However, the experiment itself, with a greater sample size, may provide a clearer picture.

After this Pilot Test, one change was implemented for Experiment 1. It was decided that the pause between the subject and the rest of the utterance in the SC stimuli would be removed because it was an additional cue of focus that was not present in the PPC.

Appendix F

Pilot Test of Flanker Task

For this task, a pilot test was conducted. A total of ten pre-schoolers (7 males and 3 females) between 4;4 and 5;11 ($\bar{x}$=4;9, *SD*=0.60) took part in the experiment and were recruited online using the social media profiles of the UNAM Laboratory of Psycholinguistics.

Before the test, the parents were given an instructional video that showed them how to glue the red and green pieces of paper to their keyboards. Then, a Zoom link was provided so they could connect and ask any questions to the experimenter before the task. The link to the platform Gorilla was then provided in the chat box. The experimenter remained on connected though videoconference with the microphone and video off to assist in case of any technical problem. This helped the experimenter to hear and see the child's reaction to the test and also to ensure the task was running correctly.

Before the analysis, data cleansing was performed. Blank responses and responses with less of 300 ms of RT (which was marked by Höhle et al., 2016 as RTs too short to give a thoughtful answer) were eliminated; this accounted for 5.2% and 1.6% of the data respectively, with a total 6.8% of eliminated data.

Regarding the analysis itself, all correct responses given were above chance (CC: 89.23%; IC: 86.87%; NC: 89.17%). According to a Binomial test performed because of the small number of participants, all conditions were significant ($p < 0.001$). These results lead to the conclusion that the test was adequate for 4-year-old pre-schoolers of this population (Höhle et al., 2016).

There was no significative differences between conditions in accuracy (CI-CC: $Z(9)= -0.638$, $p= 0.517$; CN-CC: $Z(9)= 0$, $p= 1$; CN-CI: $Z(9)= -0.962$, $p= 0.336$) or RT (CI-CC: $Z(9)= -0.866$, $p= 0.386$; CN-CC: $Z(9)= -0.561$, $p= .575$; CN-CI: $Z(9)= -0.255$, $p= 0.799$), which may be

explained by the small sample size. However, the purpose of the pilot testing itself was to prove its adequacy to test pre-schoolers of this population, a condition that was met.

Appendix G

Individual results from Flanker Task

Participants of the Flanker Task

For this analysis, all children were considered (SC, n= 20; OC, n= 21) first to obtain individual results.

Results from the Flanker Task

All correct responses given were above chance (Congruent Condition or CC: 96.29%, $t(40) = 37.23$, $p < .001$; Incongruent Condition or IC: 94.54%, $t(40) = 31.05$, $p < .001$; Neutral Condition of NC: 95.82%, $t(40) = 43.88$, $p < .001$). There was no significant difference between these scores (CC-IC: $Z(40) = 0.411$, $p = 0.681$; CC-NC: $Z(40) = 0.566$, $p = 0.571$; IC-NC: $Z(40) = 0.431$, $p = 0.667$). These results show that the task was suitable for pre-schoolers of 4 and 5 years of age.

Regarding the analysis of RTs, only correct responses were included. The average RT between condition was different: pre-schoolers were slower when they had to answer to the IC ($\bar{x}$=1597.03 ms, *SD*=688.75), followed by the CC ($\bar{x}$=1552.40 ms, *SD*=659.85) and the NC ($\bar{x}$=1402.77 ms, *SD*=561.73). These results were as expected (Höhle et al., 2016): an interference effect (the subtraction of the RTs of the IC minus the RTs of the CC) was found ($\bar{x}$=44.64 ms), in which, when presented with contradicting information (when they saw the target facing the opposite direction than the rest of the fish), children were slower to answer. Furthermore, a flanker effect (the subtraction of the RTs of the CC minus the RTs of the NC) was also found ($\bar{x}$=149.63 ms), in which a greater number of stimuli caused slower RT as seen in the CC (were multiple fish facing the same direction are presented) in comparison to the NC (were only one fish is presented). In addition, the difference of RTs between conditions was significant (CC-NC:

Z (40) = 3.557, *p*= 0.001; NC-IC: *Z* (40) = 4.257, *p*= 0.001); except for the CC and the IC (CC-IC: *Z* (40) = 1.492, *p*= 0.136) where pre-schoolers performed similarly. However, individual differences were found.

Individually, pre-schoolers ranged from having a negative interference effect (which means they had faster RTs to the IC than to the CC) as much as -716.605 ms to 588.036 ms. These individual differences were examined to determine if they had a relation with prosodic focus comprehension and will be addressed in Experiments 1 and 2.

Appendix H

Pilot Test of Experiment 2

A pilot test with adults was performed to establish the size of the time window in which an effect will be seen, before applying the experiment to pre-schoolers. For this purpose, 26 adults (16 females) of 20 to 45 years of age ($\bar{x}$= 27.30, *SD*= 6), were recruited. From this sample, 13 participants were assigned to the SC and 13 to the PPC randomly using the platform *Gorilla* (https://app.gorilla.sc/), which redirect them to one of 4 different sequences of either condition in the platform *RealEye* (https://www.realeye.io/).

Data processing

The Areas of Interest (AoI) consisted of the six images presented (an example is seen in Figure H1); all of them of an equal size of 440x340 px within a 1920x1080 px background. Two images were classified as either the subject or the prepositional predicative of the sentence (for example, considering Figure H1, “The ant is above the squirrel”, were the ant is the subject and the squirrel, the prepositional predicative), two images were classified as competitor images, as they were contrasted in the utterances presented (following the example presented before, the frog would be the subject competitor and the owl the prepositional predicative competitor, as the ant is actually above the owl and the squirrel is actually below the frog), and the last two images were classified as distractors, as they were not considered in the utterance.

Figure H1

Example of Stimuli and AoI of the Pilot Test of Experiment 2

Before the analysis, data treatment was performed. First, data was processed, where an interpolation of 150 ms was applied to exclude data loss due to blinking. In addition, to eliminate noise caused by lack of gaze precision, a Gaussian filter was applied.

Then, the proportion of data per trial was calculated, and trials with less than 20% of data on the window of analysis were eliminated, following other studies where a similar criterion was used (10% in the case of Angulo-Chavira & Arias-Trejo, 2021 and Arias-Trejo et al., 2022; 17% in the case of Quam and Swingley, 2014; and 20% in the case of Borovsky et al., 2016). A total of 810 from 936 trials remained in the SC, which accounted for 87% of the trials; while in the PPC, a total of 780 from 936 trials remained, which accounted for 83% of the trials.

After this procedure, only participants with less than 50% track loss were considered for the analysis, as done before with eye-tracking studies with children (Höhle et al., 2016; Reuter et al., 2023). For this purpose, the percentage of gazes per participant in each condition was measured. All participants remained.

To establish the size of the time window in which an effect will be seen, a first analysis was done with the total duration of each trial. Then, to compare statistical differences between conditions, the time axis was divided in two; the effect was only found in the second section: from the onset of the prepositional predicative to the end of the utterance, where comparisons between images were expected (in the SC, from 2880 to 5812 ms; and in the PPC, from 2500 to 5532 ms).

Cluster-Based Nonparametric Analysis

The first analysis consisted of a cluster-based nonparametric analysis (Maris & Oostenveld, 2007), which improves the sensitivity of the statistical tests by solving the multiple comparisons problem in data with multiple time stamps such as the eye-tracking data, in which each pair of time stamps needs to be compared. This method has been successfully used in previous eye-tracking experiments with toddlers (Mani & Von Holzen, 2012; Angulo-Chavira & Arias-Trejo, 2021; Arias-Trejo et al., 2022; Chow et al., 2022). It formulates a null hypothesis, which states the equal probability of distribution in the experimental conditions, to prove that the nonparametric test controls the rate of falsely rejecting this null hypothesis.

Each condition was analysed separately; looks to the distractors were averaged to eliminate individual effects (Höhle et al., 2016). A difference in looks is expected depending on the emphasis on the subject and the prepositional predicative in both conditions.

The data was divided in bins of 50 ms for the analysis. To compare the differences between looks to the subject competitor and the prepositional predicative competitor, *t*-tests were used. For each condition, clusters were created, were *t*-values were equal or greater than a critical value of 0.05 (used for two-tailed hypothesis). Then, the differences between looks to the subject competitor and the prepositional predicative competitor were assigned with positive and negative

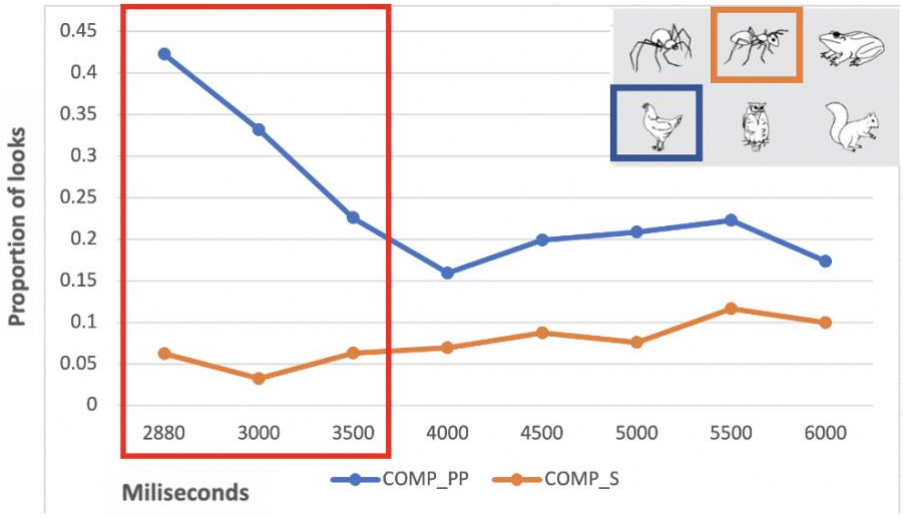
values randomly, creating 1000 iterations, and *t*-tests were applied for each iteration. This created a permuted distribution. Finally, clusters were identified; and this clusters were considered significant if their *t*-values were equal or greater than the *t*-value of the observed data. A *p* value was then obtained by dividing the number of significant clusters between the total number of iterations, which were 1000.

Before performing the analysis of the time window in which the effect was found, the total duration of the trials was analysed. For this, looks to the subject competitor and the prepositional predicative competitor were compared using *t*-tests. For each condition, clusters were created, where *t*-values were equal or greater than a critical value of 0.05 (used for two-tailed hypothesis). No significant clusters were found in either condition.

A lack of difference between conditions in the complete time window may hide significant differences in a smaller window. Because of this, a cluster analysis, as described before, was made from the onset of the prepositional predicative, in the SC, from 2880 to 5812 ms; and in the PPC, from 2500 to 5532 ms. According to this cluster analysis to the SC, the prepositional predicative competitor had a greater proportion of looks than the subject competitor from 2880 to 3500 ms (*tclust*= 11.326; tmax= 4.58; *p*=0.015). In comparison, a cluster analysis to the PPC revealed no difference in proportion of looks between the subject and the prepositional predicative competitors (*tclust*= 1.145; tmax= 1.145; *p*=1.001). These results can be seen in Figure H2 and H3:

Figure H2

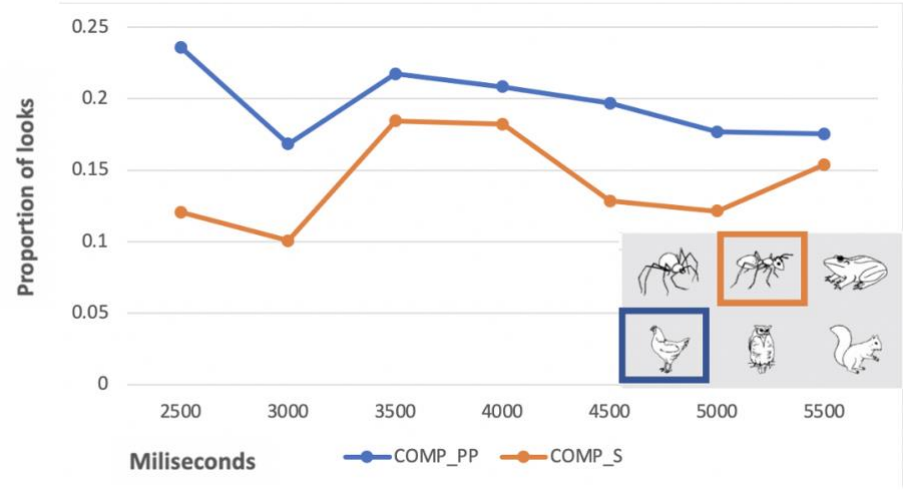
Looks in the Subject Condition



Note. COMP_PP refers to the prepositional predicative competitor; COMP_S, to the subject competitor.

Figure H3

Looks in the Prepositional Predicative Condition



Note. COMP_PP refers to the prepositional predicative competitor; COMP_S, to the subject competitor.

These results show that adult Spanish speakers had a greater proportion of looks at the prepositional predicative in the SC. This can be interpreted as the participants looking for the correct element below the subject that was emphasized. However, no further conclusions can be drawn as there were no differences between conditions in the PPC. This could be explained by the lack of context adult speakers had in this Pilot Test (as in the Pilot Test of Experiment 1); they only heard random sentences as they looked at the screen. Another explanation is the way the data was analysed.

Some changes were necessary when analysing data in Experiment 2, like the bins in which data is divided. In this Pilot test, data was divided in bins of 50 ms, which may not be the best option. In other studies, bins of 100 ms are suggested for the analysis because of the different resolutions of the participants' cameras, which might have slowed the signal (Reuter et al., 2023). This means that, in the window of analysis of each trial, data needs to be divided every 100 ms in order to make comparisons between trials.

Additionally, the complete time window still needs to be analysed as the prosodic emphasis in the SC happens at the beginning of the utterance. A story was also created for pre-schoolers, so they had to look for the correct answer as if the task was a game. These changes were implemented in Experiment 2.