The Spanish adaptation of MAIN

Maria José Ezeizabarrena

University of the Basque Country (UPV/EHU)

Isabel García del Real

Universidad Pública de Navarra (UPNA)

In this paper, we present some features of the European Spanish adaptation of the Multilingual Assessment Instrument for Narratives (LITMUS-MAIN), most of them related to specificities of the Spanish grammar as compared to English, the source language of the original MAIN (Gagarina et al., 2012). These two languages differ in e.g. 1) the use of 3rd grammatical person to address the hearer; 2) the ways of maintaining nominal cohesion: English (non-pro drop) vs. Spanish (pro-drop); 3) the verbal paradigm with regard to morphological tense and aspect morphology. Finally, preliminary results for micro- and macrostructure measures in the narratives of children with Spanish as L1 and L2 confirm their consistency across MAIN stories and procedures.

1   Introduction

The Language Impairment Testing in Multilingual Settings – Multilingual Assessment Instrument for Narratives (LITMUS-MAIN, hereafter MAIN), developed in 2012 by an international research group, is designed to assess children's comprehension and production of narratives (Gagarina et al., 2012; 2015). It includes four picture-based stories, each of them in the form of a set of six pictures, which can be used as visual support for the elicitation and/or the comprehension of their corresponding narratives. The four stories were designed with a parallel micro- and macro-structure, so that they had a very similar degree of complexity and could be used with children from three years of age. The instrument has been used to assess mostly oral narrative skills, though it may also be used for assessing participants' skills in the production and comprehension of written narratives.

Universality is one of the features of the instrument, since MAIN is intended to be universal in different ways: culturally and (psycho-)linguistically. It aims to be as culturally neutral as possible, so that it can be used to assess children's narrative skills, receptive and/or

expressive, regardless of their linguistic, social and cultural backgrounds. The large number of languages into which it has been adapted so far guarantees its linguistic universality, and this is reinforced by their genetic variability, as well as by the typological distance between many of them. The instrument is intended to be sensitive to participants' monolingual, bilingual and multilingual profiles. The instrument makes it possible to test bilingual participants in their two languages using comparable stories, which also opens the door to the possibility of testing multilinguals in more than two of their languages. Moreover, the revised version of MAIN (Gagarina et al., 2019) includes some modifications which have increased the clarity of some instructions and the easiness in the scoring.

In this paper, we present some details of the *Instrumento multilingüe para la evaluación de la narración* (IMEN), the adaptation of MAIN to Spanish. Spanish, a Romance language of the Indo-European language family, is the fourth largest language in the world, with around 463 million native speakers and around 537.9 million speakers in total (Bernhard, Simons & Fenning, 2020). More specifically, the current paper focuses on the variety of Spanish spoken in Spain. *Español* 'Spanish' or *Castellano* 'Castilian' is the official language in Spain and (one of) the first language(s) of the majority of the population. *Español* or *Castellano* is co-official with other Romance languages such as Catalan, Valencian, Aranese and Galician, and the non-Indoeuropean Basque language in the Spanish regions in which these vernacular languages are spoken. It is also co-official with Spanish Sign Language, and Catalan Sign language in Cataluña and Aragón autonomous communities. All these languages are written using the Latin alphabet.

Spanish shares many lexical roots with the other Romance languages (e.g. French, Italian), and also the (non-rigid) SVO basic constituent order, in contrast to the Latin SOV. In the nominal domain, Spanish has a very reduced case system, which is restricted to pronouns, but has overt gender and number marking in the nominal domain (nouns, determiners, adjectives and pronouns) and a very rich system of verbal inflection, where verbs are specified for person, number, tense, aspect and mood. Spanish is a pro-drop language with very frequent omission of the (lexical and pronominal) subject, though the person and number inflection of the verb identifies the grammatical person of the non-overt subject.

The frequent use of pro-drop in adult and child Spanish (Bel, 2003; Ezeizabarrena, 2013) makes the identification of the reference difficult, especially in narratives, in which most sentences contain subject-less verb forms inflected for third person singular. Another challenge in the production and comprehension of narratives in Spanish is the use of verb inflections, i.e. correct markers of aspect and tense. Studies in the acquisition of these markings have revealed that 5-year-olds understand and produce adult-like past verb inflection, but that they have difficulties interpreting and producing imperfective past verb forms in Spanish (García del Real, 2010; García del Real, van Hout & Ezeizabarrena, 2014; Garcia del Real, 2015).

In what follows, we describe some typological differences between the source language English (the language version on which all MAIN-adaptations are based) and the target language Spanish that are relevant for the adaptation of MAIN to Spanish. Moreover, we present insights based on our experience in the use of the instrument to collect and analyse data and on the preliminary results.

2   Linguistic features and their relevance in the adapted Spanish version

The MAIN text is addressed to researchers and professionals who aim to assess participants' narrative skills. It includes different levels of content that needs to be translated and adapted to the language in which it is to be used: the description of the test materials (pre-test phase), the description of the experimental procedure (test phase) and the criteria for scoring (post-test phase, scoring sheets). The typological differences between English and Spanish does not affect the adaptation of all components of the MAIN protocol equally. For instance, the differences in the ways of marking grammatical person in the two languages (Section 2.1) affects the descriptions of the materials and the procedure, and the text of the scoring sheets as well, whilst the obligatory (English) vs. optional (Spanish) presence of overt subjects (Section 2.2) and the complexity of the verb morphology (Section 2.3) affects the materials and the scoring, but not the description of the procedure.

2.1   Grammatical person

The experimental procedure section describes what the experimenter should do and say when interacting with the participant, and, consequently, this part of the text includes instructions for two different addressees, the adult experimenter, who reads the protocol and will run the test, and the test participant (child or adult), who will not read the protocol but will carry out the narrative task based on the instructions.

The Spanish grammar distinguishes three persons in the pronominal system and in the verb inflectional systems: the 1st person corresponds to the speaker (1a), the 2nd corresponds to the listener/addressee (1b, 1d), and the 3rd corresponds to non-human referent(s) or to humans which do(es) not participate in the conversation. The third person is marked for masculine (M) and feminine (F) gender. Notice that the third person (1c, 1e) was also called the "non-person" by Benveniste (1966) and is zero-marked in many unrelated languages.

(1)   a. yo   habl-o   b. tú   habla-s   c. el/ella   habla-Ø
            1s.pron   speak-1s   2s.pron   speak-2s   3s.M/F.pron speak-3s
            'I speak'   'you (singular) speak'   'he/she speaks'

        d. vosotros hablá-is   e. ellos/ellas   habla-n
            2pl.pron speak-2pl   3pl.M/F.pron   speak-3pl
            'you (plural) speak'   'they speak'

In Spanish (as in German, French and in many other languages) there are two ways to address the interlocutor: using either the informal or the formal register. In informal registers, the 2nd person morphology is used, marking pronouns and verbs for the 2nd person singular (1b), or the 2nd plural (1d), in case of more than one addressee. The formal register requires the use of the singular pronoun *usted* (2a) or the plural *ustedes* (2c) 'you singular/plural' in European Spanish, but in contrast to other languages and varieties of Spanish, these formal pronouns agree with 3rd-person-inflected verb forms (2a or 2c), instead of with the 2nd-person-inflected one (1b, 1d).

(2) a. Usted   habla-Ø   b.   Ustedes   habla-n
      2.pron.formal speak-3s   2.pl.pron.formal speak-3pl
      'You speak (singular, formal)'   'you speak (plural, formal)

In English, imperative forms are not inflected, and they are very rarely preceded by a personal pronoun, whilst the rest of finite verbs need to have a nominal or pronominal subject. In contrast, Spanish imperatives are always inflected, and the informal imperative (3a) versus formal subjunctive (3b) contrast is maintained among them.

(3) a. ¡Cuent-**a**   la   historia!   b.   ¡Cuent-**e**   la historia!
      tell-3s.Indic the story   tell-3s.Subjunctive   the story
      'tell the story! (2s informal)'   'tell the story! (2s formal)'

The optionality between two different forms of address depending on the choice of formal or informal register becomes useful to distinguish the two different addressees involved in the MAIN protocol: the experimenter (the most formal, as in (2a, 3b)) and the child participant (the less formal, as in (1b, 3a)). Thus, in contrast to the four bare infinitives in the English MAIN (4a), the Spanish MAIN distinguishes the informal inflected imperatives (cuént**a**me 'tell me') as in (3a, 4b) from the formal ones (*aním**e**le* 'encourage' (4b), *señal**e*** 'point' (4b), cuent**e** 'tell' (3b).

(4) a. …<u>encourage</u> the child to <u>tell</u> the story by him/herself by saying: *"<u>Tell</u> me the story"*
      (<u>point</u> to picture)
   b. …aním**e**le al niño a que cuent**e** la historia, diciéndole: *cuéntame el cuento* (señal**e** la imagen).

2.2   *The optionality of the subject*

Pro-drop languages allow subject (and in some cases also object) arguments to be lexical (5a), pronominal (5b) or null (5c). Spanish differs from other Romance and non-Romance pro-drop languages in the high rate of null subjects (over 60% across child and adult corpora) and the marked character of overt pronouns, including masculine (M) and feminine (F) 3rd person personal pronouns, whose use is very restricted (mostly human referents, used for contrast and focus), and consequently non-frequent. In fact, null subjects are the default option in child and adult Spanish and in many other pro-drop and non-pro drop languages as well (Bel, 2003; Ezeizabarrena, 2013). This feature has direct consequences in the production of narratives, since it directly affects the nominal cohesion and the interpretation of 3 person referents.

(5) a. el chico   estaba-Ø contento   b.   él/ella   estaba-Ø   content-o/-a
      the-M boy was-3s   happy-M   3s.pron-M/F was-3s   happy-M/F
      'the boy was happy'   'he/she was happy'

   c. estaba-Ø content-o/-a
      was-3s   happy-M/F
      'he/she was happy'

In the absence of lexical subject arguments, the identification of characters and the nominal cohesion, in general, becomes a difficult task, especially in long narratives with many characters. This may not be a major issue with MAIN, where there are three main animate characters and they are involved in actions where the thematic roles are rarely reversible: the cat/dog/fox chases the butterfly/birds/mouse/goats, the cat/dog steals the fish/sausages, the mother bird/goat protects her children (and not vice versa) and the limited number of six pictures reduces considerably the number of potential cases of referent misidentification. However, at the same time, the visual support may reduce the need for the story (re)teller to produce lexical arguments, especially in the case of young children, who may tend to use deictic and (overt or null) pronominal expressions instead, even though there is no context of shared visual attention between the child and the adult experimenter. The communicative situation may induce children to build oral texts which are closer to sequences of short descriptive utterances than to cohesive well contextualized coherent narrative texts.

Moreover, grammatical features such as number and gender may help the (partial) identification of characters in languages with inflected pronouns like English or Spanish (6b), but, even then, the identification is not always guaranteed. For instance, the masculine singular feature of the subject pronoun in *he hurt himself* is not enough to disambiguate between its two animate referents *mouse* and *dog* in the passage of the Dog story, and probably in neither language (6a, 6b). In fact, both Spanish translations, the more literal with a personal pronoun (6b), and the more natural with a null subject (6c), result in grammatical sentences, but neither option completely solves the reference problem. However, including the personal masculine pronoun *él* as the experiencer of *hacerse daño* 'hurt him/herself' and of *enfadarse* 'be(come) angry' increases the number of potential candidates, since the *él* pronoun could be associated with an additional human male referent such as the *boy*.

(6) a. The mouse ran away quickly and the dog bumped into the tree. **He** hurt **himself** and was very angry.
   b. El ratón se escapó corriendo rápidamente y el perro chocó contra el árbol. **Él se** hizo daño y se enfadó.
   c. El ratón se escapó corriendo rápidamente y el perro chocó contra el árbol. Ø **Se** hizo daño y se enfadó.

Nevertheless, there are many cases in the MAIN stories where morphological (gender and number) marking on nominal categories may reduce ambiguity by lowering the number of potential referents for both overt and null 3rd person pronouns. For instance, in (7) the feminine singular inflection of the adjective *content-a* 'happy-F.sing' excludes *gusano* 'worm.M', *crías* 'baby animal.F.pl', and *gato* 'cat.M' as potential referents for the null subject of *estaba* 'was'. In some way, the Spanish feminine ending *-a* "compensates" for the lack of the feminine pronominal subject *she,* which was sufficient to exclude the same potential referents in the original English text.

(7) a. The mother bird came back with a big worm for her children, but **she** did not see the cat. **She** was happy …

b. La madre de los pájaros regresó con un gusano enorme para sus crías, pero no Ø vio al gato. Ø Estaba muy content**a** …

2.3   The verbal system

The Spanish morphology is rich in the verbal domain, with verb forms being inflected for 1st, 2nd, and 3rd person, singular and plural number, present/past/future tense and indicative/subjunctive/potential-unreal/imperative modes. Present tense is regularly used to refer to the speech time (8a), but can also be used to refer to past events (historic present) (8b), which can be used to express perfective (8c) and imperfective past events in narrative contexts.

(8)   a.   aquí   el chico   tiene-Ø   una caña de pescar
             here   the boy   have-3s.present   a rod of fishing
             'here the boy has a fishing rod'

       b.   y entonces   el rey   dimite-Ø
             and then   the king   resign-3s.present
             'and then the King abdicated'

       c.   y entonces   el rey   dimit-ió
             and then   the king   resign-3s.past.indef
             'and then the King abdicated'

Grammatical aspect is also coded in the verbal inflectional morphology. Similarly to English, Spanish durative predicates are regularly expressed by periphrastic forms using the present/past auxiliary *estar* 'to be' followed by the imperfective participle of the lexical verb bearing the imperfective suffix -*ando* 'ing*'* (9).

(9) la cabra   se   está/estaba   ahogando
    the-F goat   3s.reflexive   is/was   drown-IPF
    'the goat is/was drowning'

Perfective (10) and imperfective (11) predicates have a different paradigm distribution in English and Spanish. In Spanish, events which are culminated in the "close" recent past (*this morning/week/year*) can be expressed by the present perfect periphrastic forms conformed by the auxiliary *haber* 'have' followed by the participle bearing the perfective suffix –*do* (10a) or by the *indefinido* 'aorist' tense (10b).

(10) a. el balón   ha caído   al río   b. el balón   cayó   al río
        the ball   has fall-Past.PF to the river   the ball   fall-Past.PF   to the river
        'the ball has fallen/fell into the river   'the ball fell into the river'

(11) el pajarito   tenía   hambre
     the baby-bird   have-Past.IPF   hunger
     'the baby bird was hungry'

Spanish pre-school children tend to tell stories in the present tense and insert past inflected verb forms as they grow older. Similarly to other adaptations (e.g. Bulgarian, see Meier & Kuehnast, 2020), we think that the use of indicative present forms in the story-telling task could be equally appropriate, as children may interpret the pictures as actions occurring in the speech time.

Acquisition studies in L1 Spanish have shown that neither children's interpretation nor production of past imperfective forms is adult-like at the age of 5 (García del Real, 2015), since they tend to interpret both perfective and imperfective telic predicates as culminated, and they tend to produce imperfective telic predicates to refer to culminated events in a context in which the use of the perfective would be more appropriate, as in (12), as this aspect conveys completion.

(12) Mientras sonaba   la música, el payaso dibujaba   una flor
    While play-Past.IPF the music, the clown draw-Past.IPF a flower
    'While the music was playing, the clown was drawing a flower" (in a context in which
    there is completion: the clown drew a flower)

Moreover, imperfective forms are used frequently in child spontaneous speech, in imaginary play contexts, sometimes overriding the distinction between perfective and imperfective forms for the reference to completion (13) as reported by Algrem and Idiazabal (2001).

(13) Ahora yo era   el médico y tú te caías y te rompías una pierna
    Now I was-IPF the doctor and you fall.IPF and break.IPF a leg
    'Now (let's imagine that) I am the doctor and you had fallen and broken a leg'

This finding is relevant for the interpretation of children's productions as referring to attempts or results, as in the one produced by a child in (14), where the use of the imperfective form may refer either to an attempt or to a result. This ambiguity is problematic in the case of incremental theme predicates, but not so in the case of accomplishments, as in (15).

(14) y el gato   estaba estaba comiendo un pez y el niño se puso   muy
    and the cat was was eat-IPF a fish and the boy SE put.PF very
    contento   porque recuperó   su pelota
    happy   because recover-PF his ball
    'and the cat **was…was… eating/ate** a fish and the boy **was** very happy because he
    recovered the ball'

(15) se   caía   su balón y no lo podía   recuperar
    3s.reflex fall-Past.IPF his ball and not it can-Past.IPF recover
    'his ball **fell** down and he **could** not **recover** it.

3 Preliminary results obtained with the Spanish MAIN-version

In this section, we report preliminary results from a pilot study based on narrative data elicited with the Spanish MAIN. The data come from 12 five-year-old children (mean age: 5;7 range:

5;1-6;1). Each of the narrative samples was collected for a different research purpose, which means that some narratives were collected in combination with the comprehension task, whilst others were collected combining the two elicitation procedures, with or without comprehension questions. Moreover, the use of the four different stories was not counterbalanced. As a result, for narratives elicited from children with Spanish as L1, the distribution over stories and tasks is as follows (Table 1):

Table 1: Number of task instances for L1 Spanish.

Story	Comprehension	Narrative production	
		Retelling	Telling
Baby Birds	7	2	5
Baby Goats	0	1	2
Cat	6	3	0
Dog	10	4	0
Total	23	10	7

With respect to the four different stories (Baby Birds, Baby Goats, Cat, Dog), the results presented in Table 2 show that there were no statistical differences between them in the pilot sample for the comprehension score ($F(2,20) = 0.828$, $p = .451$), and neither in the story structure score ($F(3,13) = 0.839$, $p = .496$) nor in the amount of internal state terms (IST) ($F(3,13) = 0.668$, $p = .587$) contained in the narratives produced.

Table 2: Mean scores for the different MAIN stories, L1 Spanish.

Story	Comprehension score		Story structure score		IST	
	Mean	SD	Mean	SD	Mean	SD
Baby Birds	8.00	1.73	7.86	2.41	4.72	3.20
Baby Goats	---	---	6.33	1.53	2.33	1.53
Cat	8.83	1.17	8.00	1.00	3.67	0.58
Dog	7.80	1.69	8.50	1.29	3.75	2.02
Total	8.13	1.57	7.29	1.86	3.88	2.39

In relation to the type of elicitation procedure, as shown in Table 3, the comprehension scores tend to be higher in the retelling than in the telling task. In contrast, story structure scores are higher, and there are more internal state terms (IST) in the telling than in the retelling task. However, none of these differences are significant (comprehension: $U = 12.5$, $p = .432$; story structure: $U = 17$, $p = .088$; IST: $U = 23.5$, $p = .270$).

Table 3: Mean scores by elicitation mode, L1 Spanish.

Task	Comprehension score		Story structure score		IST	
	Mean	SD	Mean	SD	Mean	SD
Telling	8.14	1.46	6.60	1.50	3.20	1.75
Retelling	7.40	1.37	8.29	1.97	4.86	1.22
Total	8.13	1.57	7.29	1.86	3.88	2.39

Regarding macro-structural complexity, the most frequent structure is the one that mentions the single goal (43%), and the least frequent is the one that includes the goal, the attempt and the outcome (7%). This distribution is constant for all four stories and in both tasks.

Finally, we used the Spanish MAIN with the aim of investigating whether it would be useful to distinguish between monolinguals' and bilinguals' performance. Therefore, six additional children with Spanish as an L2 were tested (mean age: 5;6, range: 4;5-6;1). However, none of the differences in mean scores shown in Table 4 are significant (comprehension: $U = 12.5$, $p = .432$; story structure: $U = 17$, $p = .088$; IST: $U = 23.5$, $p = .270$).

Table 4: Mean scores depending on the child's linguistic profile.

Linguistic profile	Comprehension score		Story structure score		IST	
	Mean	SD	Mean	SD	Mean	SD
L1 Spanish	8.13	1.57	7.29	1.96	3.88	2.39
L2 Spanish	7.25	1.84	5.83	2.12	2.75	1.81
Total	7.77	1.72	6.69	2.07	3.41	2.21

4　Conclusion

The adaptation of MAIN to Spanish and to many other languages constitutes an important step in the development of instruments for measuring mono-, bi- and multilingual children's narrative skills. Preliminary results on micro- and macrostructure measures in MAIN-narratives from L1 and L2 Spanish-speaking children confirm the consistency of the instrument across stories and elicitation procedures.

Language-specific features may pose a challenge for the accurate adaptation of MAIN as well as for a unified interpretation of the variability of responses observed across language versions. Nevertheless, an instrument adapted to many typologically distant languages will also contribute to the identification of grammatical (and lexical) development indexes. The few grammatical features discussed in this paper (person marking, null subjects, aspect and tense

inflection) are just the first in a long list of relevant linguistic features which can be considered for a promising cross-linguistic comparative research, based on the high number of languages to which MAIN has been adapted already.

5   Acknowledgments

6   References

Almgren, M., & Idiazabal, I. (2001). Past tense verb forms, discourse context and input features in bilingual and monolingual acquisition of Basque and Spanish. In J. Cenoz & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 107–130). Amsterdam: John Benjamins.

Bel, A. (2003). The syntax of subjects in the acquisition of Spanish and Catalan. *Probus*, *15*, 1–26.

Benveniste, E. (1966). *Problèmes de linguistique général I*. Paris: Gallimard.

Berhard, D. M., Simons, G.F., & Fennig, C.D. (Eds.). (2020). *Ethnologue: Languages of the World. Twenty-third edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com

Ezeizabarrena, M.J. (2013). Overt subjects in early Basque and other null subject languages. *International Journal of Bilingualism*, *17*(3), 309–336.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohacker, U., & Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, *56*.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2015). Assessment of Narrative Abilities in Bilingual Children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children disentangling bilingualism from language impairment* (pp. 243–269). Bristol: Multilingual Matters.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., & Walters, J. (2019). MAIN: Multilingual Assessment Instrument for Narratives. Revised version. *ZAS Papers in Linguistics*, *63*.

García del Real, I. (2010). La adquisición del contraste aspectual perfectivo/imperfectivo en castellano: semántica y discurso. In M.R. Caballero Rodríguez & M.J. Pinar Sanz (Eds.) *Modos y formas de comunicación humana* (pp. 163–170)*.* Ediciones de la Universidad de Castilla-La Mancha.

Garcia del Real, I. (2015). *The acquisition of Tense and Aspect in Spanish*. Unpublished PhD thesis. University of the Basque Country.

García del Real, I., van Hout, A. & Ezeizabarrena, M. J. (2014) Comprehension and production of grammatical aspect in child Spanish: Semantics vs. pragmatics. In Chu, C-Y. (Ed.) *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)* (pp. 99–110) Somerville: Cascadilla Proceedings Project.

Meier, E. & Kuehnast, M. (2020). Storytelling and retelling in Bulgarian: a contrastive perspective on the Bulgarian adaptation of MAIN. *ZAS Papers in Linguistics*, *64*, 11–21.